Journal of Industrial Information Technology and Application Vol.4 No.2

ISSN(Online): 2586-0852

Online Multiple Object Tracking using a Detection Mean Confidence-based Track Management Method

Young Chul Lim [1,], Minsung Kang [1*]

[1] Research Division of Future Automotive Technology, Daegu Gyeongbuk Institute of Science & Technology, Daegu, Republic of Korea

Abstract. In this paper, we propose a track management scheme based on the detection mean confidence to create correct tracks and to terminate false tracks efficiently and reliably. The detection mean confidence (DMC)-based tracking method follows a detection-by-tracking framework which utilizes the output of an object detector as the observation model of a Bayesian filter and updates the target state with a motion model and an observation model. In online multiple object tracking methods, detected objects are associated with current tracking objects in each frame. Accordingly, these methods are more vulnerable to noisy detections than offline methods. In this paper, we propose a reliable track management method to tackle these problems. Our experimental results show that our approach improves the accuracy of multiple object tracking while minimizing the numbers of false positives and false negatives.

Keywords; multiple object tracking; track management; object detection

1. Introduction

In recent years, object detection performance has progressed significantly due to well-established benchmarks and deep learning techniques [1, 2]. Accuracy improvement in object detectors makes it reasonable and reliable to track multiple objects such as vehicles and pedestrians despite complex environments. Many multiple object tracking (MOT) methods are based on the tracking-by-detection framework which associates and updates detections from object detectors on each frame. MOT methods are generally categorized as offline (batch) and online (sequential) approaches.

---

[*] Corresponding author: ninolyc@dgist.ac.kr

Received: Oct 18, 2019; Accepted: Dec 6, 2019; Published: Sep 30, 2020

The former method constructs trajectories of target objects by formulating a global optimization scheme using all available information over the whole video sequence [3, 4]. According to the MOT benchmark [5], the offline methods generally achieved better accuracy than the online methods due to much more information about video sequence. However, their limitation is that they cannot be used in applications where online real-time processing is required. On the other hand, the online methods [6, 7] estimate the optimal state of the object using only previous image information, and they can be applied to online applications such as real-time video surveillance systems and autonomous driving vehicles. However, the online methods are vulnerable to noisy detections and incorrect data associations lead to critical trajectory errors of target objects.

Many real-time applications such as video surveillance, intelligent robots and vehicles, are suitable for online MOT approaches which are restricted to utilizing past frames only to process information. Our MOT framework follows a real-time online paradigm, and we propose a track management method to improve both the accuracy and reliability of object detections.

The rest of the paper is organized as follows. In Section 2, we describe our confidence-based track management method in an online MOT framework. Experimental results are described and analyzed in Section 3. Finally, Section 4 provides the conclusion and future work.

2. Proposed method

A. *Online multiple object tracking method*

Our multiple object tracking framework follows the online detection-by-tracking framework which utilizes the output of object detectors as the observation model of a Bayesian filter and updates the target state with a motion model and the observation model. In the online multiple object tracking method, detected objects are associated with tracked objects in each frame. Accordingly, these methods are more vulnerable to noisy detections than offline methods. In this work, we propose a track management scheme based on the detection mean confidence to create correct tracks and to terminate incorrect tracks efficiently and reasonably. Our approach improves the accuracy of multiple object tacking while minimizing the numbers of false positives and false negatives.

B. *Detection mean confidence-based multple object tracking*

Our multiple object tracking framework is based on that in our previous work [7], which consists of multi-channel feature generation, pedestrian detection, visual tracking, and data association to increase the computational efficiency through the sharing of multi-channel features. Multi-channel features composed of three color channels and seven gradient channels are generated from a given input image. In the object detector, feature vectors are established by aggregating the multi-channel features. The visual tracker operates on the color channel and gradient channel images using a multi-channel kernelized correlation filter scheme [8]. The Hungarian algorithm-based data association task assigns multiple detections to multiple tracks by calculating the similarity costs of the histogram-based appearance model and the spatial overlapping between detections and tracks.

Unlike an earlier method [7], the proposed multiple object tracking framework consists of object detection, visual tracking, data association, and track management, and optimal target state update, as shown in Fig. 1. In this framework, the CompACT [9] is used to localize multiple class vehicles such cars, buses, vans, and others. In this paper, we focus on the track management scheme to verify the correct target object robustly. Our track management approach manages various track states, such as track inactivation, track initialization, track activation, and track termination, using the mean confidence of consecutive detection outputs and a minimum track length. A track is created from a detection which is not associated with any previous tracks. The created track is activated when the mean confidence is above a predefined threshold and the number of consecutive detections exceeds a predefined minimum track length. Moreover, a detection can be immediately activated when the detection score is very high. An activated track is in the termination state when the track misses a corresponding detection pair. The track in the termination state can be in the track activation state when the mean confidence or the detection score exceeds the corresponding predefined threshold, and the track can be in the track inactivation state when the mean confidence is below the predefined threshold. Tracks in the track activation state and the track termination state are determined as valid tracks.

Fig 1. The proposed online multiple object tracking framework

3. Experimental results

A. Experimental setting

The performance of our multiple object tracking (MOT) method was evaluated according to the rules specified in the DETRAC benchmark [10]. The dataset consists of ten hours of videos captured with a Canon EOS 550D camera at 24 different locations in Beijing and Tianjin in China. The videos are recorded at 25 frames per second (fps), with a resolution of 960×540 pixels.

Our hardware platform was a desktop computer with an Intel i7-7740X core, a working frequency of 4.30 GHz, and 32 GB of RAM. All algorithms were implemented using C++ language in Windows 10 with Visual Studio 2015. Baseline detections of CompACT [9], R-CNN [11], ACF [12], and DPM [13] were provided in the DETRACK tracking benchmark. In this work, we selected the CompACT detector [9] and evaluated the performance in comparison with previous MOT methods.

B. Experimental results

The DETRACK benchmark for object tracking consisted of 60 training sequences and 40 testing sequences. Only the vehicles outside the ignoring regions were annotated and evaluated in the benchmark. A total of 8250 vehicles were annotated in the benchmark, which consisted of 5936 vehicles in the DETRAC-train set and 2314 vehicles in the DETRAC-test set. The vehicle category had four sub-classes which corresponded to cars, buses, vans, and others. However, they were not discriminated during the evaluation. The DETRAC MOT metrics considered both object detection and object tracking. The MOT evaluated the tracking performance through the eight metrics of PR-MOTA, PR-MOTP, PR-IDS, PR-MT, PR-FRAG, PR-ML, PR-FN, and PR-FP. The PR-MOTA score was the baseline evaluation metric in the DETRAC MOT system, while the other metrics were used for reference only.

Four baseline detections for the training and test sequences were provided in the tracking challenge. In this work, we selected CompACT detection due to the superiority of detection accuracy of this approach. First, we investigated the tracking accuracy on the DETRAC training sequences according to many parameters, such as the minimum track length (Lm), the minimum mean confidence for track creation (Ccmin), the high confidence for immediate track initialization (Ch), the number of temporal windows for the detection mean confidence (Nw), and the mean confidence for track death (Cdmin). In the MOT framework, it is difficult to find optimal parameter values with a machine learning algorithm. Therefore, we manually searched for and found the optimal values by dividing and adjusting the parameters so that the MOTA had the maximum value. In this paper, we show the results of MOTA convergence after tuning several iterative parameters. In the experiment, we set the parameters Ch, Cdmi, and Nw to 0.4, 0.1, and 4, respectively. The experimental results showed that the MOT method achieved the highest MOTA score when Ccmin was 0.2 and Lm was 1 or 2 (See Table 1). In the next experiment, we evaluated the tracking performance according to changes of Nw and Ccmin while the parameters Ch, Cdmi, and Ccmin were set to 0.4, 0.1, and 0.2, respectively. As shown in Table 2, the method achieved the best MOTA score when Nw was 4 or 5. We compared our method to previous state-of-the-art methods on DETRACK test sequences. The experimental results showed that although most of the other methods were based on offline MOT approaches, our method achieved the best performance with a competitive processing speed (See Table 3).

TABLE I. MOT EVALUATION RESULTS ACCORDING TO CHANGES OF THE PARAMETERS CCMIN AND LM OF THE DETRAC TRAINING SEQUENCES

Cc_{min}	L_m	PR-MOTA	PR-MOTP	PR-MT	PR-ML	PR-IDS	PR-PRAG	PR-FP	PR-FN
0.0	1	23.5%	41.9%	**20.9%**	**16.6%**	1481.4	1827.6	21713	**140276**
	2	24.4%	41.9%	20.2%	16.9%	1161.2	1545.5	14179	142400
	3	24.6%	42.0%	19.7%	17.0%	996.4	1416.2	11879	143699
0.1	1	24.7%	41.9%	20.9%	16.7%	1364.6	1743.2	14510	140629
	2	24.9%	41.9%	20.2%	16.9%	1108.8	1494.7	11434	142553
	3	24.9%	42.0%	19.7%	17.0%	972.1	1391.8	10292	143785
0.2	1	**25.1%**	42.0%	20.2%	16.9%	1040.8	1462.9	10114	142834
	2	**25.1%**	42.0%	19.7%	17.1%	885.7	1290.6	8802	144616
	3	25.0%	42.0%	19.4%	17.2%	814.9	1231.5	8296	145369
0.3	1	24.6%	42.0%	18.8%	17.7%	780.0	1215.4	7313	148797
	2	24.5%	**42.1%**	18.4%	17.8%	704.3	1100.4	6828	149883
	3	24.2%	**42.1%**	18.3%	17.9%	**671.1**	**1065.5**	**6658**	150412

TABLE II. MOT EVALUATION RESULTS ACCORDING TO CHANGES OF PARAMETERS LM AND NW ON THE DETRAC TRAINING SEQUENCES

L_m	N_w	PR-MOTA	PR-MOTP	PR-MT	PR-ML	PR-IDS	PR-PRAG	PR-FP	PR-FN
1	1	23.6%	43.1%	18.3%	18.6%	3007.3	4245.6	7175.9	152420
	2	24.6%	42.8%	19.3%	17.7%	1697.7	2323.6	8134.3	147277
	3	25.0%	42.3%	19.9%	17.2%	1250.5	1719.4	9052.0	144511
	4	**25.1%**	42.0%	20.2%	16.9%	1040.8	1462.9	10114.8	142834
	5	**25.1%**	41.7%	**20.4%**	**16.7%**	964.8	1359.3	11108.8	**141895**
2	1	23.5%	**43.2%**	17.8%	18.8%	2739.5	3849	**6297.1**	154008
	2	24.5%	42.7%	18.8%	17.9%	1518.1	2102.4	6959.8	148861
	3	24.9%	42.3%	19.3%	17.4%	1092.8	1544.6	7785.5	146054
	4	**25.1%**	42.0%	19.7%	17.1%	885.7	1290.6	8802.7	144616
	5	**25.1%**	41.8%	19.9%	16.9%	**812.3**	**1197.3**	9778.4	143438

TABLE III. THE COMPACT DETECTOR.

Method	PR-MOTA	PR-MOTP	PR-MT	PR-ML	PR-IDS	PR-FP	PR-FN	Speed (fps)
GOG [14]	14.2%	37.0%	13.9%	19.9%	3334	32092	180183	**389.51**
CMOT [15]	12.6%	36.1%	16.1%	18.6%	285	57885	**167110**	3.79
H^2T [4]	12.4%	35.7%	14.8%	19.4%	852	51765	173899	3.02
IHTLS [16]	11.1%	36.8%	13.8%	19.9%	953	53922	180422	19.79
DCT [17]	10.8%	**37.1%**	6.7%	29.3%	**141**	13226	223578	2.19
CEM [18]	5.1%	35.2%	3.0%	35.3%	267	**12341**	260390	4.62
DMC	**17.4%**	36.6%	**15.4%**	**18.9%**	1134	20544	173605	133.57

4. Conclusions

In this paper, we proposed an online multiple pedestrian tracking method using detection mean confidence-based tracking management. The proposed method creates correct tracks and terminates incorrect tracks efficiently and reasonably based on the track state which is determined by the accumulated detection mean score and the number of associations between track and detection. Our experimental results show that our approach can achieve superior performance with a fast processing time compared with other methods. In the future, we will research a deep learning-based visual tracking method and re-identification algorithm to handle occlusion problems.

Acknowledgment

This work was supported by the DGIST R&D Program of the Ministry of Science, ICT. It was also funded by Daegu Metropolitan City and Daegu TechnoPark (Project name: Research institute cooperation convergence R & D project. 2019)

References

[1] M. Everingham, L. J. V. Gool, C. K. I.Williams, J. M.Winn, and A. Zisserman. "The pascal visual object classes VOC challenge," IJCV, vol. 88, no. 2, pp. 303-338, 2010..

[2] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll´ar, and C. L. Zitnick, “Microsoft COCO: Common objects in context,” in Proc. ECCV. pp. 740-755, 2014

[3] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally optimal greedy algorithms for tracking a variable number of objects. in Proc. IEEE CVPR, pp. 1201-1208, 2011.

[4] L. Wen, W. Li, J. Yan, Z. Lei, D. Yi, and S. Z. Li. Multiple target tracking based on undirected hierarchical relation hypergraph. in Proc. IEEE CVPR, pp. 1282-1289, 2014.

[5] A. Milan, L. Leal-Taixe, I. Reid, S. Roth, and K. Schindler, MOT16: A benchmark for multi-object tracking. arXiv preprint arXiv:1603.00831, 2016.

[6] M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller–Meire, and L. Van Gool, Online multiperson tracking-by-detection from a single, uncalibrated camera. IEEE Trans. PAMI, 33(9):1820-1833, 2011.

[7] Y. C. Lim and Ms. Kang. Multi-Pedestrian detection and tracking using unified multi-channel features. in Proc. IEEE AVSS, 2017.

[8] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. High-speed tracking with kernelized correlation filters. IEEE Trans. PAMI, 37(3):583-596, 2015.

[9] Z. Cai, M. Saberian, and N. Vasconcelos. Learning complexity-aware cascades for deep pedestrian detection. in Proc. IEEE CVPR, pp. 3361-3369, 2015.

[10] L. Wen, D. Du, Z. Cai, Z. Lei, M. Chang, H. Qi, J. Lim, M. Yang, and S. Lyu. UA-DETRAC: A new benchmark and protocol for multi-object tracking. CoRR, abs/1511.04136, 2015.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. in Proc. IEEE CVPR, pp. 580-587, 2014.

[12] P. Doll´ar, R. Appel, S. Belongie, and P. Perona. Fast feature pyramids for object detection. IEEE TPAMI, 36(8):1532-1545, 2014.

[13] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained partbased models. IEEE TPAMI, 32(9):1627-1645, 2010.

[14] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally optimal greedy algorithms for tracking a variable number of objects. in Proc. IEEE CVPR, pp. 1201-1208, 2011.

[15] S.-H. Bae and K.-J. Yoon. Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning. in Proc. IEEE CVPR, pp. 1218-1225, 2014.

[16] C. Dicle, O. I. Camps, and M. Sznaier. The way they move: Tracking multiple targets with similar appearance. in Proc. IEEE CVPR, pp. 2304-2311, 2013.

[17] A. Andriyenko, K. Schindler, and S. Roth. Discrete continuous optimization for multi-target tracking. in Proc. IEEE CVPR, pp. 1926-1933. IEEE, 2012.

[18] A. Andriyenko and K. Schindler. Multi-target tracking by continuous energy minimization. in Proc. IEEE CVPR, pp. 1265–1272. IEEE, 2011.