Multihypothesis Trajectory Analysis for Robust Visual Tracking

Dae-Youn Lee[1], Jae-Young Sim[2], and Chang-Su Kim[1]
[1]School of Electrical Engineering, Korea University. [2]School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology.

Most tracking algorithms experience a drift problem due to various reasons, including non-discriminative feature descriptors, occlusions, and sudden illumination changes. For more robust tracking, recent tracking algorithms attempt to overcome such interferences. Since a certain feature type may fail to distinguish a target object from its background, multiple trackers using different features can be combined adaptively [3, 6]. However, if the multiple-tracker algorithm loses the position of a target object in a frame because of any interruption, the tracking error may propagate to future frames. Recently, tracking systems with memory [4, 5, 8], which can refine past trajectories or appearance models of a tracker, have been proposed to suppress the error propagation. In [4, 5], most probable positions of a target object are memorized in each frame, and then the trajectory of the target object is estimated by dynamic programming. Also, in [8], several appearance models from past frames are recorded and processed to yield a proper appearance model and reduce tracking errors. However, these trackers with memory employ fixed feature descriptors, which cannot effectively separate a target object from its background in some sequences.

In this paper, we propose a novel multihypothesis tracking algorithm, referred to as the multihypothesis trajectory analysis (MTA) tracker, which combines the concept of the 'tracking using multiple trackers' with that of the 'tracking with memory.' We employ three forward trackers using different features, which are based on texture information, color information, and illumination invariant information, respectively. From frame $t-\tau$ to frame t, each forward tracker traces a target object independently of the other trackers. Then, at frame t, each backward tracker is initialized at the estimated position of the corresponding forward tracker, and then computes a backward trajectory in a time-reversed manner. To select the best tracking result among the three forward trackers, we calculate their robustness scores. To this end, we extract the geometric similarity, the cyclic weight, and the appearance similarity from each pair of the forward and backward trajectories. After selecting the best forward trajectory, the appearance models of all forward trackers revert to the previous conditions at frame $t-\tau$, and are updated using the bounding boxes along the selected forward trajectory. When all forward trajectories have low geometric and appearance similarities in consecutive frames, the forward trackers are not updated and the search range is increased for next frames. The main contributions of this work are as follows.

- Novel multihypothesis trajectory analysis to extract the best trajectory from a set of multiple trackers.
- Design of the robustness score of a pair of forward and backward trajectories, based on the geometric similarity, the cyclic weight, and the appearance similarity.
- Pattern analysis of geometric similarities and appearance similarities along trajectories to detect and handle tracking failures.

Figure 1 illustrates two trajectory hypotheses from frame $t-\tau$ to frame t. In this example, we track the target athlete within the gray bounding box at frame $t-\tau$. Red and purple curves depict the forward trajectories, obtained by two different forward trackers, and blue and green curves are the backward trajectories of the corresponding backward trackers, respectively. Note that the red tracker fails to trace the target, and the associated appearance model is continuously updated with inaccurate samples. At time t, the blue tracker is initialized at the blue bounding box. It then follows the non-target object backwardly in the reverse order of time, and the backward tracking result at time $t-\tau$ is thoroughly different from the target. In contrast, the green tracker provides the result at time $t-\tau$, which matches the original target object successfully. Therefore, we select the purple forward tracker as a valid one and discard the result of the red tracker to achieve robust object tracking.

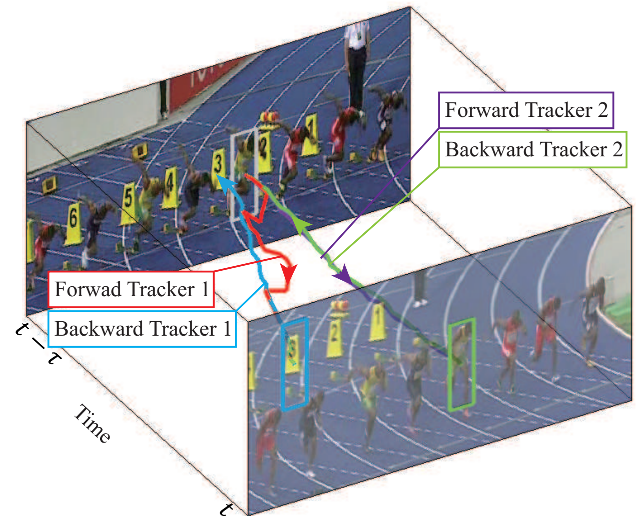

Figure 1: The forward-backward trajectory analysis: the purple forward tracker is successful, while the red one is not.

Table 1: Comparison of the average success rates (SR) and the average precision rates (PR) on the benchmark sequences in [7].

	STRUCK	KCF	MEEM	MTA
SR	0.475	0.514	<u>0.579</u>	**0.595**
PR	0.647	0.740	<u>0.836</u>	**0.838**

We test the proposed MTA algorithm on the recent benchmark dataset [7], which consists of 50 test sequences in challenging conditions, *e.g.* illumination variation, occlusion, and out-of-view. In this work, all forward and backward trackers are based on STRUCK. We compare the proposed MTA algorithm with the state-of-the-art trackers: STRUCK [1], KCF [2], and MEEM [8]. We see that MTA provides better performance than all the conventional algorithms. Especially, MTA yields 25.3% better SR and 29.5% better PR than STRUCK does.

[1] S. Hare, A. Saffari, and P.H.S. Torr. Struck: Structured output tracking with kernels. In *ICCV*, pages 263–270, 2011.

[2] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista. High-speed tracking with kernelized correlation filters. (in press) *TPAMI*, 2015.

[3] J. Kwon and K. M. Lee. Visual tracking decomposition. In *CVPR*, pages 1269–1276, 2010.

[4] Y. Lu, T. Wu, and S.-C. Zhu. Online object tracking, learning, and parsing with and-or graphs. In *CVPR*, pages 3462–3469, 2014.

[5] J.S. Supančič and D. Ramanan. Self-paced learning for long-term tracking. In *CVPR*, pages 2379–2386, 2013.

[6] F. Tang, S. Brennan, Q. Zhao, and H. Tao. Co-tracking using semi-supervised support vector machines. In *ICCV*, pages 1–8, 2007.

[7] Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In *CVPR*, pages 2411–2418, 2013.

[8] J. Zhang, S. Ma, and S. Sclaroff. MEEM: Robust tracking via multiple experts using entropy minimization. In *ECCV*, pages 188–203, 2014.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.