An Adaptive Fusion Model Based on Kalman Filtering and LSTM for Fast Tracking of Road Signs

Chengliang Wang, Xin Xie, Chao Liao

College of Computer Science- Chongqing University

Introduction

The main task of Multiple Object Tracking or Multiple Target Tracking (MOT) is to locate multiple targets of interest at the same time in a given video, maintain their IDs and record their tracks. Detection and tracking of road signs have recently gained more attention from the computer vision research and industrial community. The main objective of these algorithms is to detect and track road signs with different maintenance conditions and variable light sources from videos acquired from a moving vehicle.



Methodology

In the following Fig. 1, the proposed Fast Road-Sign Tracking algorithm (FRST) is mentioned, the FRST achieves the task by following three steps. 1) Read the t-th frame and detect the road signs through the detection model. 2) Extract visual features and motion features according to the road signs, and then assign them to trackers. 3) If the road signs assign the existing trackers, update the estimation model to estimate the state of the next frame, otherwise create a new tracker. The proposed method is described by the key components of detection, state estimation, visual features, assignment problem which is explained below in the subsections.



Fig. 1. An Illustration for the tracking of road signs from videos

Detection Model



Fig. 2. An Illustration for region growing detectors (RGD)

Estimation Model



Region growing detector (RGD)use the information about the video and the color of road signs.

For each initialized candidate X_{est} , color detection is performed on each edge.

If the ratio of color classified as a road sign **exceeds the threshold**, **it grows outward**.

Otherwise, it is deleted inward. Get a detection candidate X_{de} .



See from Fig. 3, accelerated KF has the best performance.

See from the Fig. 4, the prediction trajectory of Accelerated KF has a significant hysteresis to the true trajectory, and the prediction trajectory of LSTM is closer to true trajectory than the prediction trajectory of Accelerated KF when there are no observations.

Finally, KF model for prediction. When the confidence of the prediction is low, merge the prediction of the LSTM model.

Experimental results

Tracker	ΜΟΤΑ	Mostly tracked	False Alarm	Runti me(m s)
MOTDT [1]	92.37	91.49	4.76	97
JDE [2]	94.98	94.63	2.56	109
FairMot [3]	96.25	95.72	2.49	78
Ours(FR ST)	95.63	94.98	2.31	24

Table. 1. Comparing FRST and the state of the art algorithms

Conclusion

This paper describes an algorithm for fast detecting and tracking road signs from a sequence of video images. See from Table. 1, show FRST has the lowest running time and good performance.

References

[1] L. Chen, H. Ai, Z. Zhuang, and C. Shang, "Real-time multiple people tracking with deeply learned candidate selection and person re-identification," in 2018 IEEE International Conference on Multimedia and Expo (ICME), 2018, pp. 1-6.

[2] Z. Wang, L. Zheng, Y. Liu, and S. Wang, "Towards Real-Time Multi-Object Tracking," arXiv preprint arXiv:1909.12605, 2019.

[3] Y. Zhan, C. Wang, X. Wang, W. Zeng, and W. Liu, "A Simple Baseline for Multi-Object Tracking," arXiv preprint arXiv:2004.01888, 2020.