2018 Joint 10th International Conference on Soft Computing and Intelligent Systems and 19th International Symposium on Advanced Intelligent Systems

Tracking of Moving Vehicles with a UAV

Min-Hyuck Lee
School of Computer and Communication Eng.
Daegu University
Gyeongsan, South Korea
dool0331@daegu.ac.kr

Seokwon Yeom
School of Computer and Communication Eng.
Daegu University
Gyeongsan, South Korea
Corresponding Author: yeom@daegu.ac.kr

***Abstract*— Multiple object tracking is important for video surveillance. This paper addresses the tracking of moving vehicles with an unmanned aerial vehicle (UAV). Moving cars are segmented to generate the centroids of objects. The centroids become inputs for tracking. Tracking is performed with Kalman filtering which estimates the dynamic states of targets. In the experiment, a UAV generate video sequences of three cars driving on roads. Experimental results show that the proposed method well track the moving cars.**

Keywords*— *UAV/drone imaging, moving objects, target tracking, Kalman filtering

I. Introduction

Increased usage of unmanned aerial vehicles (UAVs) or drones becomes more popular for aerial surveillance [1]. Researches on moving object detection are conducted in literature. Background subtraction and frame difference were studied in [2–4]. Long-range moving objects were detected by background subtraction in [5].

Kalman filtering provides a solution for estimating the state of the target in real time. It is known to be optimal under independent Gaussian noise assumption [6]. When multiple measurements are detected in a frame, data association is required to assign the measurements to the established tracks [7,8]. Target tracking with UAVs has been researched in [9-12]. Fuzzy-aided Kalman filtering is proposed in [13].

In this paper, we address the tracking of multiple moving vehicles with a UAV. The UAV generates video sequences of moving objects. We integrate vision-based detection and state-based tracking. The detection is based on image processing, which manipulates pixels in the stationary scene. The tracking is based on Kalman filtering, which estimates dynamic states (position and velocity) of targets in the time domain.

In the experiments, three moving cars are captured from a long distance by a UAV. The camera built in a UAV points directly downward to the road. Experimental results show that three targets are well detected and tracked by the proposed method.

The remains of the paper are organized as follows. Object detection is presented in Section 2. Target tracking is discussed in Section 3. Section 4 demonstrates experimental results. The conclusion follows in Section 5.

II. Object Detection

Object detection is performed through frame differencing and thresholding, morphological filtering, and removing false alarms based on the minimum size of vehicles [11,12]. Two frames separated by a constant interval are subtracted and turned into a binary image by thresholding, which can represent the difference between the current and the past frames. Two morphological operations [14], erosion and dilation, are applied to the binary image to extract regions of interest (ROIs), which are the disjoint alternative areas. The ROIs are the candidate areas of moving vehicles. False ROIs are removed by the minimum ROI size, which is calculated from the true size of moving vehicles. The centroids of ROIs represent the measured positions of targets, which will be processed in the next tracking stage.

III. Target Tracking

A. System Modeling

The dynamic state of the target is modeled as a nearly constant velocity (NCV) model; the target's maneuvering is modeled by the uncertainty of the process noise, which is assumed to follow the Gaussian distribution. The following is the discrete state equation of a target:

$$\mathbf{x}_t(k+1) = F(\Delta)\mathbf{x}_t(k) + q(\Delta)\mathbf{v}(k), \quad t=1,...,N_T, \tag{1}$$

$$F(\Delta) = \begin{bmatrix} 1 & \Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad q(\Delta) = \begin{bmatrix} \Delta^2/2 & 0 \\ \Delta & 0 \\ 0 & \Delta^2/2 \\ 0 & \Delta \end{bmatrix}, \tag{2}$$

where $\mathbf{x}_t(k) = [x_t(k)\, \dot{x}_t(k)\; y_t(k)\; \dot{y}_t(k)]^T$ is the state vector at frame k, $x_t(k)$ and $y_t(k)$ are the positions in x and y directions, respectively, $\dot{x}_t(k)$ and $\dot{y}_t(k)$ are the velocities in x and y directions, respectively, T denotes matrix transposition, Δ is the sampling time, and $\mathbf{v}(k)$ is a process noise vector, which is Gaussian white noise with the covariance matrix $Q = diag([\sigma_x^2\ \sigma_y^2])$, and N_T is the number of targets. A measurement vector is composed of two components in x and y directions. The following is the measurement equation:

978-1-5386-2633-7/18/$31.00 ©2018 IEEE
DOI 10.1109/SCIS-ISIS.2018.00154

$$\mathbf{z}_t(k) = \begin{bmatrix} z_{tx}(k) \\ z_{ty}(k) \end{bmatrix} = H\mathbf{x}_t(k) + \mathbf{w}(k), \quad (3)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (4)$$

where $\mathbf{w}(k)$ is a measurement noise vector, which is Gaussian white noise with the covariance matrix $R = diag([r_x^2 \ r_y^2])$.

B. Prediction and Filter Gain

The state and covariance predictions are iteratively computed as:

$$\hat{\mathbf{x}}_t(k \mid k-1) = F\hat{\mathbf{x}}_t(k-1 \mid k-1), \quad (5)$$

$$P_t(k \mid k-1) = FP_t(k-1 \mid k-1)F^T + Q, \quad (6)$$

where $\hat{\mathbf{x}}_t(k \mid k-1)$ and $P_t(k \mid k-1)$, respectively, are the state and the covariance prediction for target t at frame k. The residual covariance $S_t(k)$ and the filter gain $W_t(k)$, respectively, are obtained as:

$$S_t(k) = HP_t(k \mid k-1)H^T + R, \quad (7)$$

$$W_t(k) = P_t(k \mid k-1)H^T S_t(k)^{-1}. \quad (8)$$

C. Data Association

Data association is the process of assigning multiple measurements to established tracks. Measurement gating is a preprocess that reduces the number of candidate measurements for data association. The measurement gating is performed by chi-square hypothesis testing with the assumption of the Gaussian measurement residuals [7]. Measurements become valid for target t at frame k if they exist in the validation region as

$$Z_t(k) = \{\mathbf{z}_m(k) \mid \mathbf{v}_{tm}(k)^T [S_t(k)]^{-1} \mathbf{v}_{tm}(k) \le \gamma, m = 1, \ldots, m_t(k)\}, \quad (9)$$

$$\mathbf{v}_{tm}(k) = \mathbf{z}_m(k) - H\mathbf{x}_t(k \mid k-1), \quad (10)$$

where $z_m(k)$ is the mth measurement vector at frame k, γ is the gating size, and $m_t(k)$ is the number of candidate measurements for target t at frame k. The nearest neighbor rule associates a measurement with track t as:

$$\hat{m}_t = \underset{m=1,\ldots,m_t(k)}{\arg\min} \| \mathbf{v}_{tm}(k)^T [S_t(k)]^{-1} \mathbf{v}_{tm}(k) \|. \quad (11)$$

D. State Estimate and Covariance Update

The state estimate and the covariance matrix of targets are updated as follows:

$$\hat{\mathbf{x}}_t(k \mid k) = \hat{\mathbf{x}}_t(k \mid k-1) + W_t(k)\mathbf{v}_{t\hat{m}_t}(k), \quad (12)$$

$$P_t(k \mid k) = P_t(k \mid k-1) - W_t(k)S_t(k)W_t(k)^T. \quad (13)$$

If $m_t(k)$ is equal to zero, i.e., no measurement is associated with target t at frame k, they merely become the predictions of the state and the covariance as:

$$\hat{\mathbf{x}}_t(k \mid k) = \hat{\mathbf{x}}_t(k \mid k-1), \quad (14)$$

$$P_t(k \mid k) = P_t(k \mid k-1). \quad (15)$$

IV. Experimental Results

A. Scenario Description

A drone (Phantom 4 Advanced) captures three moving vehicles at 30 frames per second. A total of 372 frames are captured for more than 12 seconds. The size of one frame is 4096 × 2160 pixels; the frame is reduced to 20% size, 820 × 432 pixels, for efficient image processing. The drone hovers while capturing video sequences at a height of 92 meters above the road. One pixel corresponds to 0.1344 meter.

Three cars moving, with different initial positions and maneuverings. The first car (Car 1) is located at the center of the scene from the first frame and moves to the right until it disappear after the 155th frame. The second car (Car 2) moves to the left from the first frame to the 289th frame. The third car (Car 3) moves in the same direction as the second car from the 74th to the 372nd frame. Fig. 1(a) shows Cars 1 and 2 at the first frame, Fig. 1(b) shows Cars 1–3 at the 150th frame, and Fig. 1(c) shows Car 3 at the 300th frame. In the figures, the red circles indicate the moving vehicles.

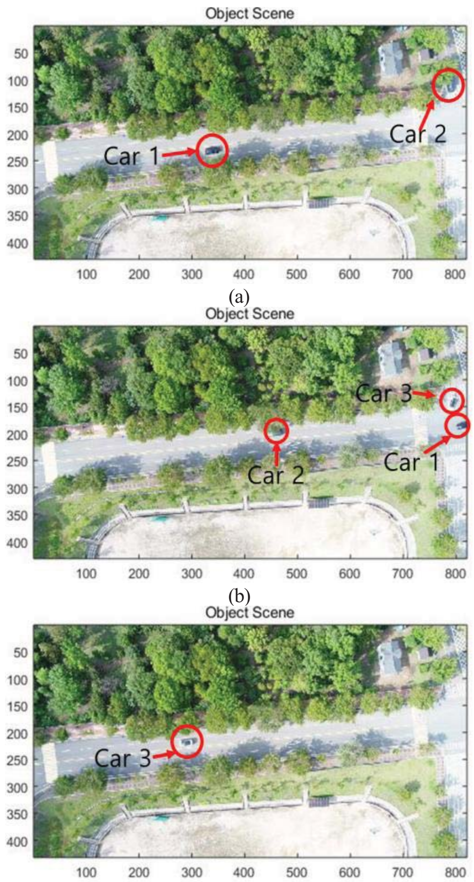

Fig. 1. (a) Cars 1 and 2 at the first frame, (b) Cars 1–3 at the 150th frame, (c) Car 3 at the 300th frame. Moving cars are indicated by the red circles.

B. Multiple Target Tracking

The centroids in Fig. 2 become input measurements for tracking. The sampling time Δ in Equation (2) is 0.033 sec since the frame rate is 30 fps. The standard deviation of process and the measurement noise are set at $\sigma_x = \sigma_y = 30\ \text{m/s}^2$ and $r_x = r_y = 1.5\ \text{m}$, respectively; γ for the gating test is set at 4. The tracks are initialized by a two-point differencing initialization method with a maximum speed gating [12]. The validation region in Equation (9) is chi-square distributed with the degree of freedom is equal to the dimension of measurement vectors, thus the probability mass corresponds to 86.5% when γ is set at 4. Fig. 3 is the results of tracking Cars 1–3.

Fig. 2. Centroids including false alarms for 368 frames.

V. Conclusion

In this paper, multiple moving objects are captured over long distances by a UAV. Three targets are tracked with Kalman filtering. Experimental results show that the proposed algorithm detects and tracks moving objects with good accuracy. This study also proposes new applications for drones, such as traffic control and smart car as well as security and defense. Tracking a large number of targets with high maneuvering remains for future study.

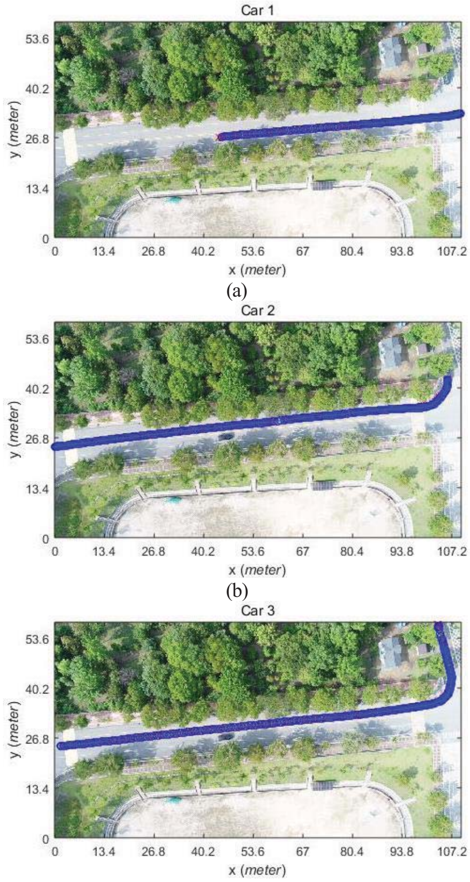

Fig. 3. Tracking results: (a) Car 1, (b) Car 2, (c) Car 3.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant Number: 2017R1D1A3B03031668).

References

[1] R. Kumar, et al., "Aerial Video Surveillance and Exploitation" IEEE Proceedings of the IEEE, vol. 89, no. 10, pp. 1518-1539, 2001.

[2] C.-H. Yeh, C.-Y. Lin, H.-E. Lai, K. Muchtar and M.-T. Sun, "Three-pronged Compensation and Hysteresis Thresholding for Moving Object Detection in Real-Time Video Surveillance," IEEE Trans. on Industrial Electronics, vol. 64, no. 6, pp. 4945-4955, 2017.

[3] Y. Wu, X. He and T. N, "Moving Object Detection with a Freely Moving Camera via Background Motion Subtraction," IEEE Trans. on Circuits and Systems for Video Technology, vol. 27, no. 2, pp. 236-248, 2017.

[4] S.-M. Park, J. Park, H.-B. Kim and K.-B. Sim, "Specified Object Tracking Problem in an Environment of Multiple Moving Objects," International Journal of Fuzzy Logic and Intelligent Systems, vol. 11, no. 2, pp. 118-123, 2011.

[5] S. Yeom, M.-H. Lee and I.-J. Cho, "Long-range Moving Object Detection based on Background Subtraction," ISIS2017, pp.1082-1085, 2017.

[6] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell, Bayesian Multiple Target Tracking 2nd, Artech house, Boston, 2014.

[7] Y. Bar-Shalom and X. R. Li, Multitarget-multisensor Tracking: Principles and Techniques, CT, YBS publishing, 1995.

[8] S. Yeom, T. Kiruba and Y. Bar-Shalom, "Track Segment Association, Fine-step IMM, and Initialization with Doppler for Improved Track Performance," IEEE Transactions on Aerospace and Electronic Systems, vol. 40, no. 1, pp. 293-309, 2004.

[9] K. Haag, S. Dotenco and F. Gallwitz, "Correlation Filter Based Visual Trackers for Person Pursuit using a Low-Cost Quadrotor," Innovations for Community Services (I4CS), 15th International Conference on, 2015.

[10] P. Chen, Y. Dang, R. Liang, W. Zhu and X. He, "Real-Time Object Tracking on a Drone with Multi-Inertial Sensing Data," IEEE Trans. on Intelligent Transportation Systems, vol. 19, no. 1, pp.131-139, 2018.

[11] M.-H. Lee and S. Yeom, "Vehicle Tracking based on Kalman Filtering with Drone Imaging," JIFS, *in press*.

[12] M.-H. Lee and S. Yeom, "Detection and Tracking of Multiple Moving Vehicles with a UAV," IJFIS, *under review*.

[13] K. Rawan-Hamid, A. Talukder, A. K. M. Ehtesanul Islam, "Implementation of Fuzzy Aided Kalman Filter for Tracking a Moving Object in Two-Dimensional Space," IJFIS, vol. 18, no. 2, pp. 85-96, 2018.

[14] R. C. Gonzalez and R. E. Woods, Digital Image Processing 3rd, Pearson Prentice Hall, New Jersey, 2008.