MULTIPLE-TARGET TRACKING FOR CROSSROAD TRAFFIC UTILIZING MODIFIED PROBABILISTIC DATA ASSOCIATION

Hsu Yung Cheng, Jenq Neng Hwang

Department of Electrical Engineering, Box 352500
University of Washington, Seattle, WA, USA

ABSTRACT

A multiple-target tracking system aimed at analyzing crossroad traffic systematically is proposed in this paper. The proposed mechanism is based on Kalman filtering and modified probabilistic data association. Unlike traditional Kalman filtering tracking, the proposed mechanism constructs candidate measurement lists by matching the sizes of the measurements and the targets first. When the sizes do not match, object matching within a limited area is performed. Also, we modify the classical Probabilistic Data Association method to enhance its performance and make it more suitable for vision-based systems. The proposed mechanism, which can serve as the foundation for automatic traffic event detection, can solve the occlusion problems effectively without incurring too much computational complexity.

Index Terms— video signal processing, crossroad traffic analysis, tracking, intelligent systems.

1. INTRODUCTION

Vehicle tracking and event detection are important applications in Intelligent Transportation Systems (ITS). Principal events include traffic congestion, dropping items, accidents, and regulation violation. It is desired to develop an automatic monitoring system that can keep track of each vehicle, analyze its behavior, recognize and report relevant events. The foundation stone of such a system is a robust tracking system that can identify each individual moving vehicle in spite of occlusion caused by camera angle. Earlier, many researchers worked on tracking vehicles and getting traffic parameters on highways, where vehicles move in mainly one principle direction [1]. Recently, the need for systems capable of handling crossroad traffic drew attentions of researchers to investigate more sophisticated systems that take multi-directional traffic into account [2], [3], instead of considering only uni-directional traffic.

Fig. 1 illustrates the system framework of the proposed system. First, a background image can be estimated from the image sequence [4], [5]. Then, based on the background image, the moving objects are separated from the background via the moving object segmentation procedure.

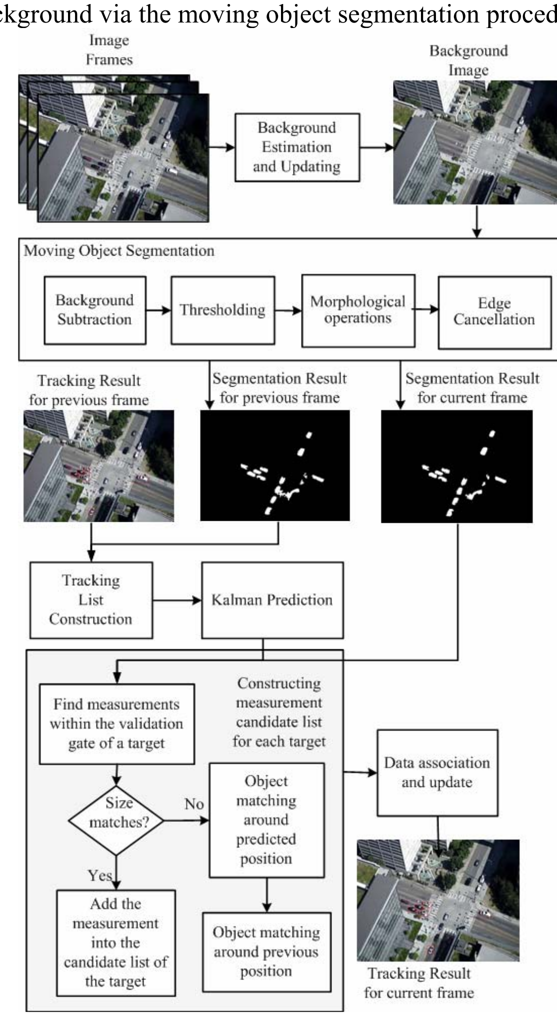


Fig. 1. System framework.

The moving object segmentation procedure is elaborated in Section 2. After that, a tracking list is constructed according to the specification of the region of interest (ROI), as explained in Section 3. For each target in the tracking list, we continue to predict its future position and velocity.

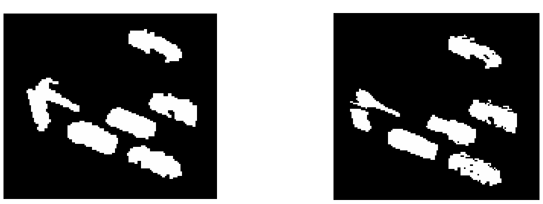
Kalman filter [6], [7] can be used for the prediction purpose, utilizing the current state plus the difference between the previously predicted state and the current measurement. Measurements are provided by the segmentation results. Because moving object segmentation may not always be reliable, we construct a measurement candidate list for every target in the tracking list and use a modified version of Probabilistic Data Association [10] to associate the most probable measurement to the target. The Kalman prediction, candidate measurement list construction and modified Probabilistic Data Association are elaborated in Sections 3, 4 and 5, respectively. Experimental results are illustrated and explained in Section 6. Finally, conclusions and future works are discussed in Section 7.

2. MOVING OBJECT SEGMENTATION

To segment the moving objects (vehicles), each image frame is subtracted from the background image to obtain the difference image. The fourth order moment of each pixel in the background difference image is calculated using the following equation [8]:

$$\mu_d^{(4)}(x,y) = \frac{1}{N_\eta} \sum_{(s,t)\in\eta(x,y)} (diff\_img(s,t) - \hat{m}_d)^4 \qquad (1)$$

where $\eta(x,y)$ denotes a window of size N_η. Each pixel (x,y) in the temporal difference image is thresholded based on $\mu_d^{(4)}(x,y)$. Afterwards, some morphological operations (opening and closing) are performed to get rid of noises. To separate different objects connected with each other and reduce the occurrence of initial occlusion, we also perform edge cancellation on the thresholded images. More specifically, the edges with the orientations parallel to the roads are detected from the original image frames. Then the pixels corresponding to these edges are set to zero in the thresholded images. Fig. 2 demonstrates the effect of edge cancellation on a thresholded image.



(a) (b)

Fig. 2. Edge cancellation: (a) Before (b) After.

3. TRACKING LIST CONSTRUCTION

We need to set up different Regions of Interest (ROI) for effective video analysis. Two examples of ROI are illustrated in Fig. 3. The selection of ROI depends on the coverage and the clarity of the video-camera. We develop a convenient graphical user interface to allow users to draw a polygon to specify the ROI. When detecting an object entering the ROI, we analyze its shape and orientation. If the shape and orientation is impossible to be a vehicle, we discard this object; otherwise we issue a vehicle ID to it and put it into the tracking list.

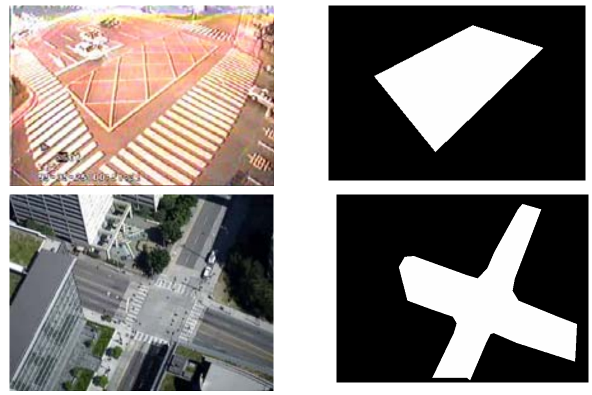


Fig. 3. Region of Interest.

4. KALMAN PREDICTION

For each object in the tracking list that has a vehicle ID, Kalman filtering is applied to predict its position in the next frame. The details of Kalman filtering are not reviewed here because they can be found in [6], [7]. In Kalman filtering, we need to define a representation point for each object. Initially, we define the center of a connected component as the representation point of an object. To apply Kalman filter, we also need to design the system state x_k, the measurement vector y_k, the transition matrix F_k, and the observation model H_k. Here, x_k and y_k are defined as $[u_k\ v_k\ \dot{u}_k\ \dot{v}_k]^T$, where u_k and v_k are the coordinates of the representation point in the image plane at time *k*. $\dot{u}_k$ and $\dot{v}_k$ are the displacements in the *u* and *v* directions, respectively. Finally, F_k and H_k are defined as:

$$F_k = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \qquad (2)$$

5. MEASUREMENT CANDIDATE LIST CONSTRUCTION

Selecting representation points for measurements is crucial to the accuracy of the tracking algorithm. Although initially we can use the centroid of a connected component as the representation point, simply relying on the centroid is not a reliable way to represent each object in the video scene. When the objects are occluded from each other due to the vision angle of the camera, different objects are connected with each other in the segmentation result. In this situation, one connected component may consist of several different objects. Therefore, it is necessary to establish a more reliable mechanism to determine the representation points for the measurements to be associated with the targets.

For each target, a validation gate is formed. y_k is in the validation gate of x_k if the following criterion is satisfied: $[y_k - H_k x_k]^T S_k^{-1}[y_k - H_k x_k] \le \gamma^2$ [9],[10]. We check each measurement within the validation gate of a target to see if it is of similar size with the target, as illustrated in Fig. 1. If the sizes of the measurement and the target match, then the measurement is put into the measurement candidate list. Otherwise, we perform object matching around the predicted position p_1 and the previous position p_2. That is, we shift the object around p_1 and p_2 to compute the overlapping area and the correlation between the segmentation result and the shifted object. If either the overlapping area or the correlation of the matching result is higher than a threshold at a particular position, then the position is treated as a representation point and added to the measurement list.

6. DATA ASSOCIATION

Because this is a multi-target environment, we need to associate one measurement with each listed target. To accomplish this, we propose a modified version of Probabilistic Data Association (PDA) which was originally proposed to track multiple targets for sonar or radar systems in cluttered environments [9] - [11]. With a slight modification, PDA could be very useful for vision-based tracking systems as well. To associate the correct measurement with the target, PDA considers each target independently, and computes a probability β_j for each measurement in the measurement list of a target. Suppose that there are in total m measurements for a given target at time k. Then the posterior association probabilities β_j can be expressed by the following equation [11]:

$$\begin{aligned}\beta_j &= P\{\chi_j \mid Y^k\} = P\{\chi_j \mid \tilde{y}_1, \cdots, \tilde{y}_m, m, Y^{k-1}\} \\ &= \frac{p(\tilde{y}_1, \cdots, \tilde{y}_m \mid \chi_j, m, Y^{k-1}) P\{\chi_j \mid m, Y^{k-1}\}}{p(\tilde{y}_1, \cdots, \tilde{y}_m \mid m, Y^{k-1})} \\ &= A_j \cdot B_j / C\end{aligned} \tag{3}$$

where χ_j denotes the event that the j^{th} measurement is the true measurement for the target. $y_1, \cdots, y_m$ denote the m candidate measurements at time k. Y^k denotes the set of all data vectors y_j, i.e., $Y^k = \{y_1, \cdots y_m\} \bigcup Y^{k-1}$. $\tilde{y}_j$ denotes the corresponding innovation for $j=1,\ldots,m$. C is the sum of the numerators. The expressions for B_j are listed as below in order to explain the modification we made on the classical PDA. The mathematical expressions and the detailed explanations for other terms can be found in [9] - [11], and therefore are not elaborated here.

$$\begin{aligned}B_j &= P\{\chi_j \mid m, Y^{k-1}\} = P\{\chi_j \mid m\} \\ &= P\{\chi_j \mid m^F = m-1, m\} P\{m^F = m-1 \mid m\} \\ &\quad + P\{\chi_j \mid m^F = m, m\} P\{m^F = m \mid m\} \\ &= \begin{cases} (1/m) P\{m^F = m-1 \mid m\} & j = 1, \cdots, m \\ P\{m^F = m \mid m\} & j = 0 \end{cases}\end{aligned} \tag{4}$$

In Equation (4), it is assumed that $P\{\chi_j \mid m^F = m-1, m\} = 1/m$ is equal for every measurement. However, in vision-based systems, we can obtain some extra information by comparing the similarity and the overlapping area between the measurement and the target. Therefore Equation (4) can be modified as below.

$$B_j = \begin{cases} \delta_j \cdot P\{m^F = m-1 \mid m\} & j = 1, \cdots, m \\ P\{m^F = m \mid m\} & j = 0 \end{cases} \tag{5}$$

where $\delta_j = \alpha \dfrac{Similarity_j}{\sum_{i=1}^{m} Similarity_i} + (1-\alpha) \dfrac{OverlapArea_j}{\sum_{i=1}^{m} OverlapArea_i}$

and α is an adjusting factor between 0 and 1. By incorporating the factor δ_j, which is more relevant to the relation between the targets and the measurements, we can obtain the probability for the modified PDA. Other terms remain the same as the classical PDA. The update phase of the covariance matrix and state estimate also remains intact.

7. EXPERIMENTAL RESULTS

Several experimental results are illustrated and explained in this section. There were in total four experimental videos used. Among the 107 vehicles, 103 are tracked correctly. The overall accuracy rate is 96.26%. Fig. 4 displays some vehicles in occlusion conditions, the corresponding segmentation results and the tracking results. Fig. 5 demonstrates the process of object matching for two occluded vehicles. The matching process is launched because the measurement size is much larger than the original size of the target. The shaded area is the measurement chosen by the modified PDA for the smaller vehicle. If the centroid of the connected component is selected as the representation point of the measurement, the result would be wrong. The proposed algorithm selects the center of the shaded area as the new representation point, which is the desired result. The corresponding tracking results for Fig. 5 are shown in Fig. 6. We can observe from these examples that the tracking system still works well when occlusion occurs. In addition, because the correlation is computed only when occlusion or segmentation error occurs and the searching is conducted within a very small range, the proposed mechanism can solve the occlusion problems effectively without incurring too much computational complexity.

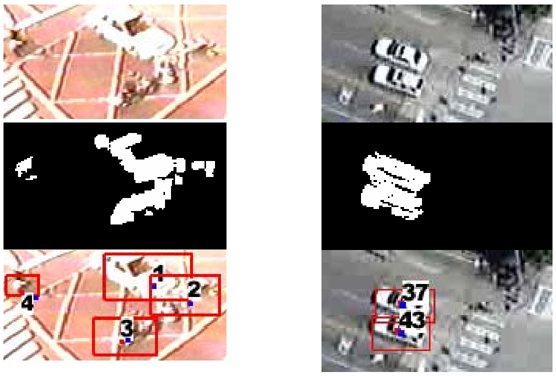


Fig. 4. Vehicles in occlusion conditions, the corresponding segmentation results and tracking results.

Fig. 5. Object matching process for two occluded vehicles.

Fig. 6. Tracking Results for Fig. 5.

8. CONCLUSIONS AND FUTURE WORKS

A multiple-target tracking system aimed at handling crossroad traffic is proposed in this paper. The proposed mechanism utilizes Kalman filtering to predict the states of the targets and employs modified PDA to associate the correct measurement with each target. To make the tracking system more robust when handling multi-directional traffic and occluded vehicles, the proposed mechanism constructs candidate measurement lists by matching the sizes of the measurements and the targets first. When the sizes do not match, object matching within a limited area is performed. Also, the classical PDA method is modified in this paper to enhance its performance and make it more suitable for vision-based systems. The proposed system can serve as a critical foundation for event detection systems at crossroads. Our future work is to develop an event detection system that can analyze the behaviors of the vehicles, recognize and report abnormal or rule-violation events in real time.

11. REFERENCES

[1] A. Goldsmith, G. W. Cleven, “Highway electronic systems - today and tomorrow,” *IEEE Trans. on Vehicular Technology*, vol. 19, no. 1, pp. 161-167, Feb. 1970.

[2] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, “Traffic monitoring and accident detection at intersections,” *IEEE Trans. on Intell. Transport. Syst.*, vol. 1, no. 2, pp. 108-118, Jun. 2000.

[3] D. Contea, et al, “A graph-based, multi-resolution algorithm for tracking objects in presence of occlusions,” *Pattern Recognition*, vol. 39, no. 4, pp. 562-572, Apr. 2006.

[4] D. R. Magee, “Tracking multiple vehicles using foreground, background and motion models”, *Image and Vision Computing*, vol. 22, no. 2, pp. 143-155, Feb. 2004.

[5] L. Li, W. Huang, I. Y.-H. Gu, Q. Tian, “Statistical modeling of complex backgrounds for foreground object detection,” *IEEE Trans. on Image Processing*, vol. 13, no. 11, pp. 1459-1472, Nov 2004.

[6] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.

[7] R. E. Kalman, R. S. Bucy, "New Results in Linear Filtering and Prediction Theory", *Transactions of the ASME - Journal of Basic Engineering*, vol. 83, pp. 95-107, 1961.

[8] S. Colonnese, A. Neri, G. Russo, G. Scarano, “Spatially adaptive HOS-based motion detection for video sequence segmentation,” *Proc. of SPIE*, vol. 5150, pp. 2015-2025, 2003.

[9] T. E. Fortmann and S. Baron, “Problems in multi-target sonar tracking,” *Proc. 1978 IEEE Conf. on Decision and Control*, Jan. 1979.

[10] Y. Bar-Shalom and E. Tse, “Tracking in a cluttered environment with probabilistic data association,” *Automatica*, vol. 11, pp. 451-460, Sept. 1975.

[11] T. E. Fortmann, Y. Bar-shalom, M. Scheffe, “Sonar tracking of multiple targets using joint probabilistic data association,” *IEEE Journal of Oceanic Engineering*, vol. OE-8, no. 3, pp. 173-184, Jul. 1983.