An Image Processing Based Motion Tracker/Estimator for Traffic Control Purposes

Ivana Vujanović, Abdalgalil Sagir Abdulla and Stevica Graovac

***Abstract*— Object tracking based on video processing is used for numerous traffic control applications. The Kalman filter has shown as a powerful apparatus in tracking of vehicles on the road. This paper is oriented to the specific tracking situations where the proper estimation of vehicle's motion has to be used due to temporarily occlusion of tracked objects. The contribution of a paper consists in a specification of proper criteria related to the switching between measurements and predictions. The algorithm was tested in some of typical road scenarios.**

***Index Terms*— Traffic surveillance, Vehicle tracking, Kalman filter, Motion estimation, Occlusion of objects.**

I. Introduction

One of the significant applications of video-based systems is the visual surveillance. Traffic surveillance system can be used to detect and track the moving objects in order to calculate the average speed of each individual vehicle, to classify vehicles, to count them in order to estimate the traffic density, to predict the congestion, etc. The extensive survey of work and problems associated with visual surveillance and object's motion and behavior is given in [1], the one focused to the moving vehicles, in [2], while the early applications in traffic context existed more than a half of century ago [3].

Videos are the sequences of images produced by TV/IR camera, each of which is called a frame. The contents of two consecutive frames are closely related. This fact is of particular use in detection of motion of objects from frame to frame. In general, the process of visual surveillance in dynamic scenes includes the following steps:

1. ***Detect*** the moving objects of interest in actual frame,
2. ***Track*** the detected objects from frame to frame (detect again and recognize each of them),
3. ***Predict*** the position of tracked objects in the next frame,
4. ***Update*** the overall number of objects according to the actual appearing/dismissing in the field of view.

The complexity of object tracking in traffic surveillance applications is mainly due to the noise in images, scene illumination changes, complex objects' motion, and occlusion.

The goal of this paper is to discuss the problem of occlusion in traffic surveillance systems and to implement the robust algorithm capable to track multiple objects continuously and reliably, by solving this issue. As a key part of the algorithm we use well known Kalman filter as an optimal filter in the sense that it minimizes the estimated state error covariance. Based on the assumed object's motion model, the Kalman filter optimally filters out the measurements related to object's position in the image and predicts its new position.

This algorithm has two main stages: target detection and target-track association.

Detection phase is here realized on the principles given in [4] based on static scene subtraction. Some specific problems related to the shadow of the vehicle, discussed in [5,6], have been analyzed also and the approach based on the segmentation according to the range of gray levels is tested too.

Tracking phase consisting in subsequent detections and track associations have been realized based on the application of a Kalman filter [7], implemented as a software module in [8], based on principles given in [9]. A number of authors characterize the tracked object via set of features related to the object's corners [10,11], while in some approaches [12], the basic feature was the appearance of windshield reflection. Our choice on this step was to use simple features as the size and position of a centroid of the tracked gate around the moving object.

Occlusion as a basic problem is analyzed in a number of papers (an extensive review is given in [13]). We applied here a simple criterion related to abrupt changes in a number of tracked objects and the size of tracking windows, as the information that the occlusion has occurred and that the switching to prediction mode should be made in the cases of occluded objects.

Section II of paper briefly discusses the analyzed ways of detection of objects of interest. Section III is representing the implementation of a suggested multiple objects tracking algorithm, where the objects are represented via their centroid and size of tracking window as the features used to describe them. The results of usage of Kalman filter/predictor in the context of temporarily occlusions of objects is given in Section IV, while some of concluding remarks are given in Section V.

Ivana Vujanović is PhD student on the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: vujanovic.iva@gmail.com).

Abdalgalil Sagir Abdulla is PhD student on the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: abdalgalilsagir@yahoo.com).

Stevica Graovac is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: graovac@etf.rs)

II. Initial vehicle detection

First phase of traffic surveillance system consists in detecting of the moving vehicles. Vehicle detection has been performed here using these methodologies:

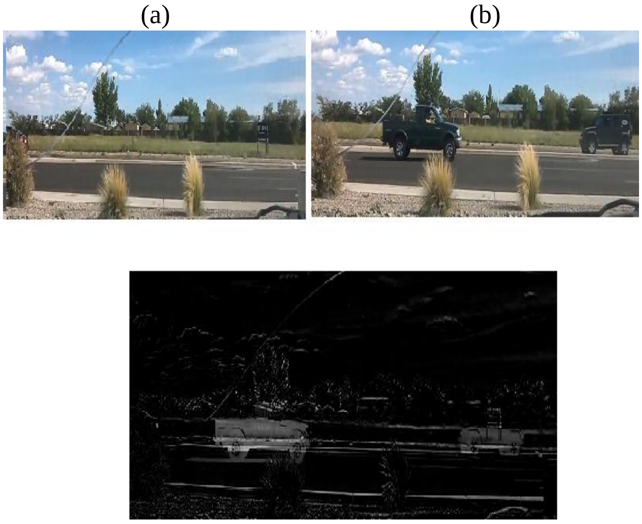
- motion of detection based on background substraction,
- detection of objects using contracts-based methods.

Both of these methods are characterized by their own advantages and disadvantages. The first one is based on the assumption that the changes in image light intensity distribution are due to object's motion only. This assumption is compromised by luminance changes (during the day/night, due to changing of weather conditions, shadows, etc.). The second method is based on the assumption that the vehicle on the road, as an object of interest, is characterized by significant difference in light intensity relative to the local background. There are a number of road scenes where it is very difficult to specify the criterion of distinguishing the object from the background and where a number of image pixels can be erroneously classified as belonging to some moving object.

A. Background substraction method

The process of extracting moving foreground objects from stored static background image is known as background subtraction [2,3]. This method involves building a representation of the scene known as the background model and finding deviations from the model for each incoming frame in video. Any significant change in an image region from the background model is noted as a moving object. In order to find the vehicles, some threshold value must be defined. Threshold is used in order to classify the image pixels as belonging to foreground or background. This process is simple and easy to realize, and it accurately extracts the characteristics of target data, but the drawback is in their sensitivity to the changes of external environment. The illustration of this principle is given in Figure 1. Two moving cars on the road (1.b) are distinguished by simple subtraction of static scene (1.a), and the difference image (1.c) clearly shows the objects of interest.

(a) (b)



(c)

Fig. 1. Usage of background subtraction.

B. Contrast-based methods

An alternative approach is to use the information about the range of light intensities characterizing the area of the object that is to be detected. This approach is mainly inoperable if one considers the whole image (for example, on the Figure 3.b it is practically impossible to distinguish the cars this way). The practical way of implementation is similar to the previous one, but it consists in specifying of some windows where the object is expected. In the most of traffic surveillance applications, these areas are close to the borders of image. Depending on the fact what is the position of camera (on the road, above the road, along the road) it is possible to specify what part of image is the candidate where one can expect the appearance of a moving vehicle. By measuring the absolute value of difference between the actual light intensities ($I_t(x,y)$) and the average gray levels inside the window (*avg(x,y)*), and by comparing it with the specificed threshold, *Diff*:

$$\left|\operatorname{avg}(x,y) - I_t(x,y)\right| > Diff\,(x,y). \qquad (1)$$

one can extract the pixels belonging to the incoming vehicle and to form the rectangular window around the “target area”. This area will grow until the vehicle is completely inside the field of view and it will be the subject of tracking that follows.

The serious problem related to both of these methods of early detection consists in the fact that the vehicle's shadow is usually classified as a moving object also and this would cause the further inter-object occlusion problems (windows around the object are sometimes unnecessarily large). In some of the examples discussed in Section IV these effects will be evident.

Whatever is the method of detection of moving vehicle, from this phase on, the subject of tracking is represented as a group of pixels characterized by two features: the size of the window around and the position of window’s centroid. There are a number of object's features that are used in different tracking applications, but in the case of real time application such as traffic surveillance, one should use simple and still relatively robust ones.

III. Multiple vehicle detection and tracking

Tracking is the process of locating the interested object within a given sequence of frames, from its appearance to its last one. A tracking system should be able to predict the position of all tracked objects, assuming that some of them are occluded. In this section, the Kalman filter algorithm in the form that is used in MATLAB in Computer Vision Toolbox is explained. The Kalman filter is an optimal recursive data processing algorithm and is a linear minimum variance error estimation algorithm for the state sequence of a dynamic system. It adopts dynamic state equation and observation equation to describe the tracking process.

In order to predict the next position of the object, we need to make an object’s motion model corresponding to the physical characteristics of its motion. In this particular example it is assumed that the vehicles on the planar road (reference coordinate frame) are moving linearly, with constant translational velocity.

$$\frac{d}{dt}\vec{x}_R = \frac{d}{dt}\begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \\ 0 \end{bmatrix} \quad (2)$$

Applying the simple formulas of perspective geometry, one can transform these trajectories onto the trajectory of object's in camera fixed coordinate frame:

$$\vec{x}_C = T_C \vec{x}_R \quad (3)$$

where T_C stands for direction cosine matrix relating camera and reference coordinate frames.

Projections onto the focal plane (image coordinate frame) are specified as:

$$x_i = f\frac{y_C}{x_C}, \quad y_i = f\frac{z_C}{x_C} \quad (4)$$

The observation equation is specified as:

$$\vec{y} = f\begin{bmatrix} \frac{y_C}{x_C} \\ \frac{z_C}{x_C} \end{bmatrix} \quad (5)$$

The motion of these points in image is going to be practically still linear, whereas the angular position of camera relative to the road would affect onto the velocity of objects' motion through the image.

From the selected frame, any object can be picked for tracking by setting the position of the mask and then object can be tracked in subsequent frames. Initial steps of algorithm include the following:

- video to frame conversion,
- preprocessing,
- background frame initialization,
- segmentation,
- morphological filtering.

As is typical for a surveillance system, the recording camera is stationary, on a bridge or beside the road, covering a large field of view. The initialization of the tracker is performed by differencing between a continuously updated background image and a newly acquired image. This is known as temporal differencing, where the background is changing along time. Morphological operations of opening and closing are applied to the resulting image. After the moving objects have been segmented, some process preparations for the subsequent tracking is needed. For the scope of this research, the arithmetic filtering technique was chosen for its simplicity and effectiveness. Then, tracking is performed by predicting the object's position from the previous information and verifying the existence of the object at the predicted position.

The aim of an object tracker is to generate a trajectory of the path followed by the moving objects over time by locating their positions in every frame of the video. Kalman filter is capable of tracking multiple objects simultaneously handling association of detection with the corresponding tracks. Tracking can be divided into the following steps:

- detecting of object
- associating the detection with existing tracks
- creating the new tracks
- deleting the object

After the object's detection, the new track is either updated on each scan, or deleted, if it has been invisible for too many consecutive frames. In the same time, where detection and tracks are unassigned, the new track is initialized. Track is confirmed if it is spotted in certain number of adjacent frames. The convenient set of track life stages is given in Table I.

TABLE I
TRACK LIFE STAGE DEFINITION

Track type	**Definition**
Potential	Single point track
Tentative	A grouping of points for which it is not assured to be object
Confirmed	A grouping of points associated to an object
Deleted	A grouping of points that became invisible for too many frames

Kalman filtering is an efficient solution for tracking of multiple objects. However, mistakes become more frequent as the number of objects increases.

IV. DEFINITION OF CRITERIA

Once the object has been detected then it can be tracked along its path. Kalman filter recursively estimates the state of the target object, having the main role when the car is lost or sheltered. In the video sequences, many of objects disappear from the scene passing behind another object with respect to a camera. Because of that, partial or total occlusions are caused. Occlusion is a significant problem in moving object detection and tracking. Generally, occlusion occurs as:

- self-occlusion,
- inter-object occlusion,
- occlusion by background scene structure

Self-occlusion most frequently arises during the tracking specified object when one part of the object occludes another. Inter-object occlusion occurs when two tracked objects occlude each other. Similarly, occlusion by background occurs when a structure in the background occludes the tracked objects. In this paper, the Kalman filter (KF) is used for handling occlusion by modeling the object motion using linear dynamic models. The type of occlusion which is our area of interest is inter-object occlusion where a vehicle merges with others. Occlusion detection is performed by detecting the overlapping regions. Overlapping areas are then specifically analyzed. In order to solve this problem we introduce an upgrade of the KF tracking algorithm. The idea is to detect the presence of occlusion and to keep the trajectory of occluded object(s) based on predictions only, until the object of interest appears in the scene separately again. In the case of tracking of multiple objects, we have to analyze the object's trajectory in order to realize where it starts with disappearing.

There are two elements of occlusion detection criterion that have been applied (Fig 2):

- The number of tracked objects is abruptly reduced (and this is not due to the moving of some object completely outside the scene);
- There was detected an abrupt change in centroid position in the new frame (greater than some threshold, either in x or y directions) due to the fact that partial occlusion will generate an artificial shift in object trajectory.

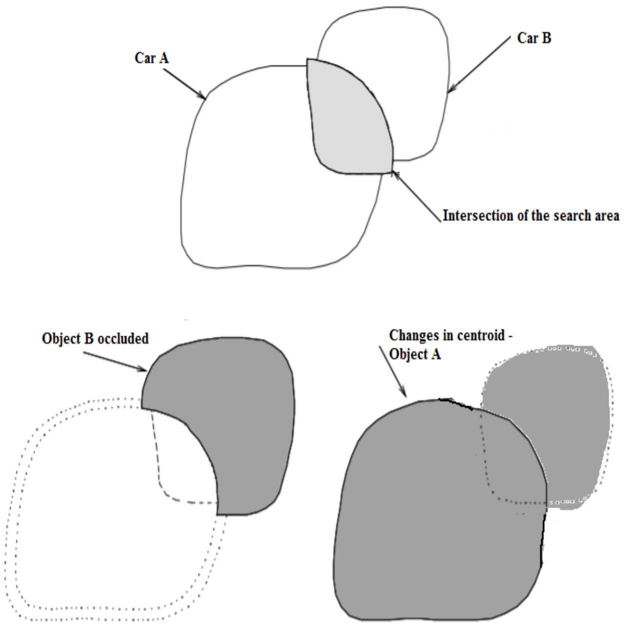


Fig. 2. Example where the occlusion occurs; The main criteria in starting the prediction in Kalman filtering.

The criterion for restarting the regular tracking based on measurements is the inverse version of previous one - after the occlusion was finished, the number of tracked objects rises sharply, while the position of centroid of a false object undergoes sudden changes again.

V. Examples of Usage

The following are the examples to illustrate the application of this algorithm.

Case Nº1 - camera above the road - inter-object occlusion

Figure 3. illustrates the scene where the camera is located above the road. The vehicles are entering and leaving the scene from top and bottom of image. Fig. 3.a represents the multi-car scene without any occlusions, while Fig. 3.b represents one of the frames where two cars formed one "false object" Fig. 3.c represents the history of positions of all cars in the scene. One can recognize easily the moment when two cars are lost and one "false car" is started with tracking. Both inter-occluded continued to be tracked via predictions, until the occlusion has been finished.

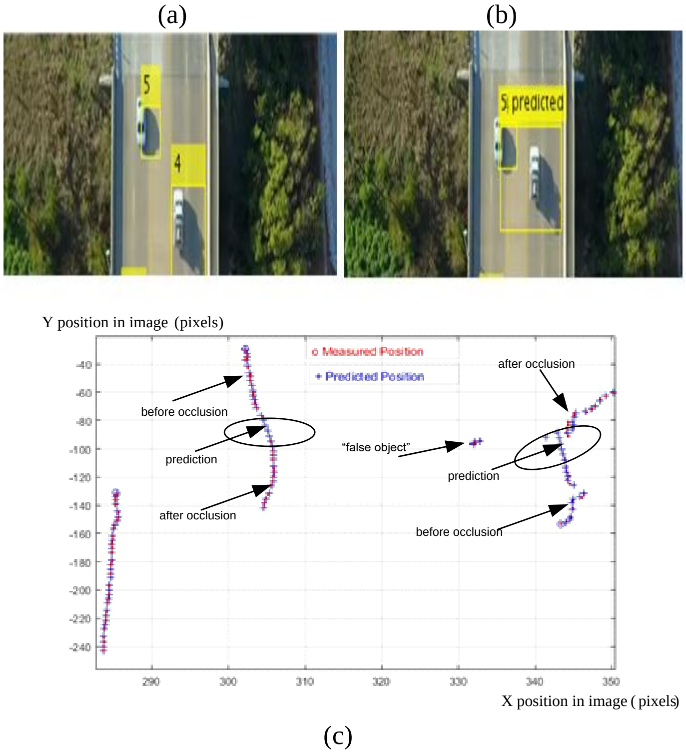


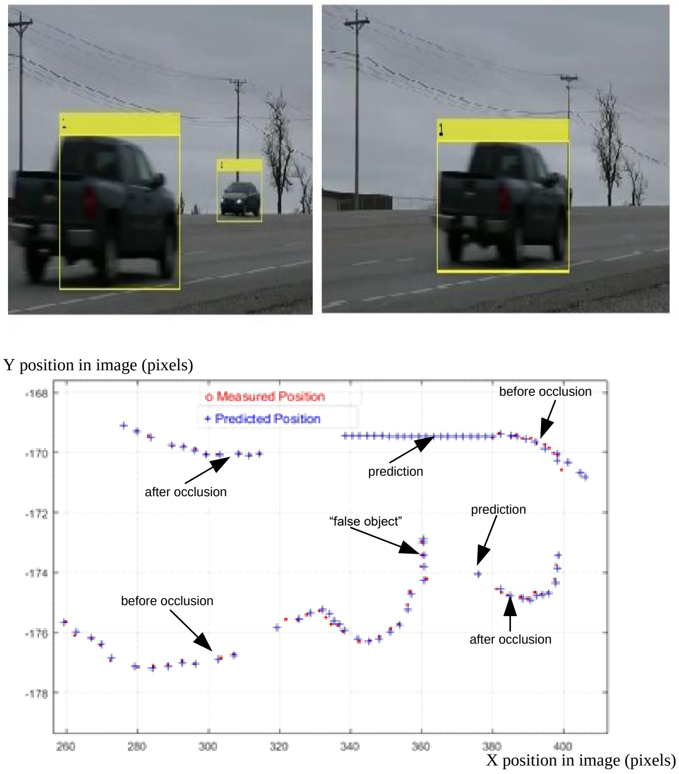
Fig. 3. Trajectories in the Case Nº1

This case is an example where the occlusion is introduced due to the vehicles' shadows included into the tracked windows. Even after the occlusion was finished, measurements related to the car on the right side are strongly affected by the car's shadow ("occlusion by background scene structure").

Case Nº2 - camera along the road - complete occlusion

Camera is located on the side of road (Figure 4.) Two cars are separately moving in opposite directions (Fig. 4.a) until the moment when one of them starts to occlude the other and in one moment the occlusion is complete. (Fig. 4.b). Trajectory on Fig. 4.c shows the fact that the outgoing car has been tracked all the time (with some abrupt changes of centroid position related to the moments of start and stop of occluding). The incoming (sheltered) car is lost in one moment and after that the information about it is obtained just based on predictions. After this car is separated again, one can see the "jump" in centroid position due to change in angular velocity existing during the interval of occlusion.

(a) (b)



(c)

Fig. 4. Trajectories in the Case Nº2

VI. Conclusion

In this paper, we proposed a method for dealing with the occlusion problem in traffic surveillance system. The implementation is made using MATLAB - Computer Vision Toolbox. Occlusion is very likely to occur in reality and the methods for object tracking generally fail in these situations. The main contribution consists in a solution to the inter-object occlusion. A proposed tracking algorithm detects when occlusion has occurred and solve the problem by holding onto the prediction based on previous motion. The solution is based on starting the prediction phase until the object of interest appears separately in the scene again. The results are acceptable even in the case of complete occlusion. Object tracking of multiple moving objects has been successfully implemented on arbitrary video sequences using Kalman Filter. Experiments encompassed different scenes and objects.

The natural extension of this research is oriented toward the situations where the occlusion is generated by the elements of the road, recognizing the cases of "inter-shadow" overlapping, and improvements of a motion model in order to cope with the problems introduced by variable angular velocity of object inside the field of view.

References

[1] H. Weiming, T. T. L. Wang, S. Maybank, "A Survey on Visual Surveillance of Object Motion and Behaviors," *IEEE Transactions on systems, man, and cybernetics-applications and reviews*, vol. *34*, no. *3*, pp. *334-352*, August, 2004.

[2] R. A. Hadi , G. Sulong, L. E. George, "Vehicle Detection and Tracking Techniques: A Concise Review", *Signal & Image Processing : An International Journal (SIPIJ) Vol.5, No.1*, February 2014

[3] J. Munkres, "Algorithms for the Assignment and Transportation Problems," *Joural of the Society for Industrial and Applied Mathematics*, vol. *5*, no. *1*, pp. *32-38*, March, 1957.

[4] S. H. Shaikh, K. Saeed, N. Chaiki, "Moving Object Detection Using Background Substraction," in *Moving Object Detection Using Background Substraction*, Cham, Country: CH of Springer International Publisher August, 2014, ch. *3*, sec. *2*, pp. *15–23*.

[5] P. Kaewtrakulpong, R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection," in *Video-Based Surveillance Systems*, New York, USA: of Springer US, 2002, ch.*11*, sec. 2, pp. *135*–144.

[6] C. Stauffer, W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," *IEEE Computer Society Conference on, Fort Collins, USA, vol. 2*, pp. 246-252, June, 1999.

[7] G. Welch, G.Bishop "An Introduction to the Kalman Filter," *UNC, CH, USA, Rep. TR 95-041*, 2006.

[8] MATLAB, Computer Vision Toolbox, Motion-based Multiple Object Tracking*:* *https://www.mathworks.com/help/vision/examples/motion-based-multiple-object-tracking.html*

[9] M. L. Miller, H. S. Stone, I. J. Cox, "Optimizing Murty's ranked assignment method,"*IEEE Transactions on Aerospace and cybernetics-Electronic Systems*, vol. *33*, no. *3*, pp. *851-862*, July, 1997.

[10] B. Coifman, D. Beymer, P. McLauchlan, J. Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance," *Transportation Research Part C: Emerging Technologies*, vol. *6*, no. *4*, pp. *271-288*, August, 1998.

[11] Z. Qiu, D. Kao "Kalman Filtering Used in Video-Based Traffic Monitoring System," *Journal of Intelligent Transporation Systems*, vol. *10*, no. *1*, pp. *15-21*, August, 2006.

[12] H. V. Pham, B.R. Lee, "Front-view car detection and counting with occlusion in dense traffic flow," *AInternational Journal of Control, Automation and Systems*, vol. *13*, no. *15*, pp. *1150-1160*, October, 2015.

[13] H. Chandel, S. Vatta, "Occlusion Detection and Handling: A Review," *International Journal of Computer Application*, vol. *120*, no. *10*, pp. *33-37*, June, 2015.