International Research Journal of Applied and Basic Sciences
© 2013 Available online at www.irjabs.com
ISSN 2251-838X / Vol, 4 (12):3658-3663
Science Explorer Publications

CONSISTENT VEHICLES TRACKING BY USING A COOPERATIVE DISTRIBUTED VIDEO SURVEILLANCESYSTEM

Peyman Babaei[*1] , Mahmood Fathy[2], Reza Berangi[2]

1. Department Of Computer, Science And Research Branch, Islamic Azad University, Tehran, Iran
2. Computer Faculty, University Of Science And Technology, Tehran, Iran

***Corresponding Author email:**p_babaei@ieee.org

ABSTRACT: This paper proposes a method of consistent vehicles tracking using a distributed cooperative multi-camera surveillance system.Vehicles' matching in views of different cameras is a fundamental issue for increasing the accuracy of consistent tracking. Among the most important research in intelligent transportation systems (ITS), automatically urban traffic surveillance is one of the critical and challenging tasks. A single-camera system has inherent limitations such as limited field of view and errors in background subtraction. The proposed work addresses synchronizing the cameras for tracking vehicles simultaneously in overlapping fields of view. Multiple cameras with overlapping fields of view can provide fault tolerance and robustness for issues such as vehicles occlusion. Distributed fundamental purpose is to efficiently reduce the transmission rate and also analyze a traffic scene and report statics and information of interests. In this manner, the surveillance area is wider and using multiple views of cameras to produce uniform tracking can enhance the capability and performance of consistent tracking. By presenting anew transmission protocol in a robust manner, we evaluate the tracking accuracy. By this way, in occlusion situations, the vehiclesin overlapping fields of view are tracked more accurately.

Keywords:Distributed Multi-Camera Systems, Overlapping FieldsOf View, Consistent Tracking.

INTRODUCTION

Multi-camera systems become increasingly attractive in machine vision Applications include multi-view vehicles tracking, behavior and event detection, occlusion handling and etc. For many applications, there may be constraints of transmission bandwidth and complexity in analyzing a huge amount of data centrally. In intelligent transportation systems (ITS), the convenient conditions are achieved by using autonomous agents making decisions in a decentralized manner. In this paper, we develop a method for tracking and recognition by a traffic video surveillance system of two distributed cameras with an overlapping field of views where processing and decision is distributed across the cameras. The number of cameras and complexity of surveillance systems have been increasing to make better coverage and accuracy. Also, multiple cameras with overlapping field of views can enhance the capability of vision application and providing fault tolerance.

In urban traffics, a surveillance system does not require prior knowledge about the events that should be recognized. The event models are automatically built by the system itself in an unsupervised way. A commonly used approach in this case is based on the clustering the trajectories of the detected moving objects, then obtained clusters are used as a normality model for abnormality detection. An abnormalevent is simply different from themost common patterns. Vehicle tracking is a challenging problem, especially in the presence of occlusion. But in this paper, we focus on the stage of vehicles tracking and analysis in a cooperative distributed video surveillance framework in urban traffic management (Mathew and Krishna, 2007).

This paper is organized as follows: an overview of the previous researches is presented in section 2. Our proposed tracking algorithm is presented in section 3. Experimental results of performance measurements with respect to Ground-truth are presented in section 4. Finally the conclusion is presented.

PreviousResearches On Multi-Camera Tracking Systems

Previous researches have been proposed to implement multi-camera surveillance systems. Multiple cameras with overlapping fields of view can enhance the capability and performance of tracking and providing robustness for vehicles occlusion. When a vehicle is occluded in one camera view, other cameras may have a better view of that vehicle. One approach is fully calibrating the cameras with considering all cameras parameters. This work is done with projection equations using coordinates of a set of points and their projections (Kelly et al., 1995). By using kalman filters and sends the features to a central server, objects are tracked. Another research assumes that only the neighboring cameras are calibrated (Cai and Aggarwal, 1999). Objects are tracked by a single camera until the system switches to another camera for a better view. Although, there are multiple cameras in the system but objects are tracked on one camera at a time.

Bayesian networks were used to match features and establish correspondences (Chang and Gong, 2001; Utsumi et al., 2000). However, relying on features matching can cause problems as the features can be seen differently by different cameras due to lighting variations. One of the feasible approaches that can be used to solve the consistent labeling is estimating a homograph (Lee, 2000). By this way, the scene is planar and a homograph is estimated between the planes in different camera views. Consistent labeling can be established with using field of view lines (Khan and Shah, 2003). This method does not require the cameras to be calibrated.

Tracking Algorithm In Overlapping Field Of Views

Vehicles trajectories may not always be tracked reliably on one camera view due to the errors resulting from the vehicle detection by the performance of the background subtraction algorithm. The problem of multi-view objects tracking has been addressedin many papers (Liu et al., 2013; Berclaz et al., 2009; Butt and Collins, 2013; Fleuret et al., 2011; Babaei and Fathy, 2010; Babaei and Fathy, 2011), but almost the surveillance system is designed centrally. In our surveillance system the position of cameras is shown in Figure 1.The system isdecentralized andeach of the cameras has four processing cores in different levels (figure 2). First, input video sequences are fed to detection level of cameras. At the Decision level, control commands are issued to classify the detected vehicles based on extracted description featuresProcess cores in three upper levels exchange the requisite information to track and recognize more accurately.

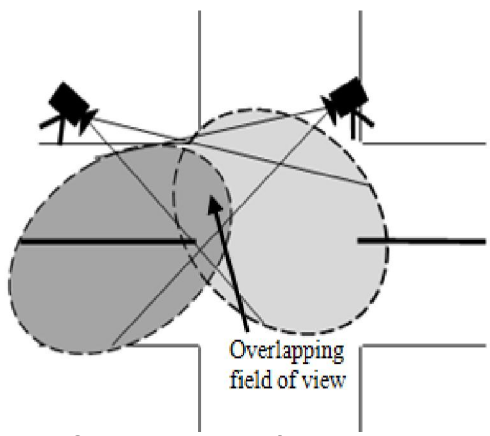

Figure1. Cameras position of cooperative system

The principle features of our scheme which is summarized as cooperative distributed video surveillance system are particularly well-suited for low bandwidth; therefore the requirement processing is done locally and the method does not require the pre-calibration into the scene and hence, can be used in traffic scenes where the system administrator may not have control over the activities taking place.

If $\mathrm{Cam} = \{\mathrm{cam1}, \mathrm{cam2}\}$be a set of two partially overlapping synchronized cameras, v_n^i represent the i^{th} vehicle observed in $\mathrm{cam}(n)$, the$T_n^i(x^i(t), y^i(t))$ is the resulting Point tracking of v_n^i on location (x, y) at instant (t) in $\mathrm{cam}(n)$. The trajectory of v_n^i is a set of all observations, where represent the length of the trajectory:

$$T_n^i = \{T_n^i(x^i(0), y^i(0)), T_n^i(x^i(1), y^i(1)), \ldots, T_n^i(x^i(L), y^i(L))\}$$

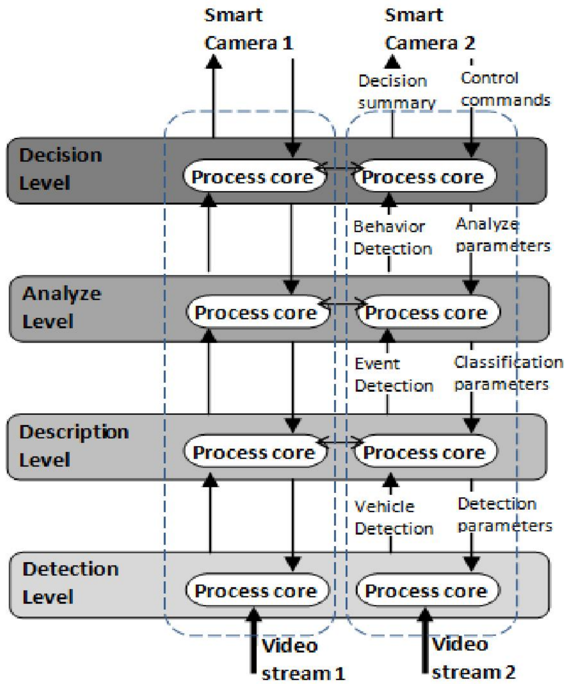

Figure2. Cooperative Levels in proposed distributed system

In order to determine the mathematical relationship between the observations from two different views, we calibrate the two cameras using five control points as shown in figure 3. In traffic surveillance, an accurate camera calibration simply means that the 2D image coordinate can be properly predicted given the 3D location of the vehicle. (Babaei and Fathy, 2011).

Figure3. Five Control Points in overlapping fields of view

The transformation from the real world position (x_w, y_w, z_w) to the 3D coordinates (x_{3D}, y_{3D}, z_{3D}) is given by a rotation matrix and a translation vector as:

$$\begin{bmatrix} x_{3D} \\ y_{3D} \\ Z_{3D} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{21} & T_{31} \\ T_{12} & T_{22} & T_{32} \\ T_{13} & T_{23} & T_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} T_{41} \\ T_{42} \\ T_{43} \end{bmatrix} \quad (1)$$

The transformation from 3D camera coordinates (x_{3D}, y_{3D}, z_{3D}) to the 2D coordinates (Q_x, Q_y) is obtained using perspective projection as:

$$Q_X = \frac{x_{3D}}{z_{3D}} f \quad , \quad Q_y = \frac{y_{3D}}{z_{3D}} f \qquad (2)$$

Using the calibration technique presented by Leibe et al., (2007) and Barrois et al., (2009), we know the homogeneous transformation T, focal length f_T and the image center $(T_x, T_y)$for camera I and S, $f_s$and $(S_x, S_y)$for camera II. The image position (Q_{1x}, Q_{1y}),$(Q_{2x}, Q_{2y})$and world position (x_w, y_w, z_w) are calculated by using the translation equations. Knowing the camera's calibration, we are able to merge the object tracking from two different views into a real world coordinates view. The final step is to fuse a pair ofcorresponding trajectories in overlapping fields as shown in figure 4. To fuse T_1^i andT_2^i, where both trajectories are generated from the same vehicle in world coordinates, we use an adaptive weighting method,

$$T^i_{1,2} = \begin{cases} \alpha T^i_1(t) + \beta T^i_2(t) & : R_{1,2} \\ T^i_1(t) & : R_1 \\ T^i_2(t) & : R_2 \end{cases} \quad (3)$$

Where $R_{1,2}$ is the region where observations from both $T^i_{1,W}$ and $T^i_{2,W}$ are available at time't'. R_1 and R_2 are the regions where the observation is available from either $T^i_{1,W}$ or $T^i_{2,W}$, respectively. The weights α and β are calculated by number of observations for each trajectory:

$$\alpha = \frac{|T^i_1|}{|T^i_1| + |T^i_2|} \quad , \beta = \frac{|T^i_2|}{|T^i_1| + |T^i_2|} \quad (4)$$

Subject to$\alpha + \beta = 1$

Finally, to construct a complete trajectory across the entire path, we connect all partial trajectories that belong to same vehicle.

$$T^i_W = \cup T^i_{1,2} \quad (5)$$

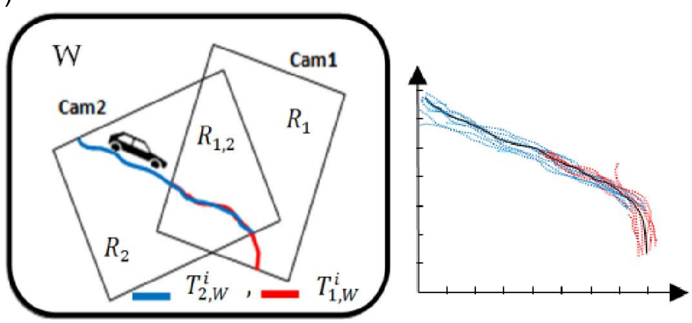

Figure4. Uniform trajectory extraction with two overlapping fields of view cameras

The global trajectory of a vehicle is obtained from connecting the two partial trajectories in world coordinates. The reconstruction of vehicles trajectories across cameras facilitates the recognition of global behaviors for large scale events in traffic monitoring and video surveillance.

Block diagram in figure 5 illustrates the implementation of the proposed cooperative protocol with communication between cameras. Both cameras will need information from the other camera when a new vehicle appears in its field of view or a vehicle tracker cannot be matched to its target vehicle. Therefore, Request messages with format illustrated in table 1 is created to be sent to other camera.

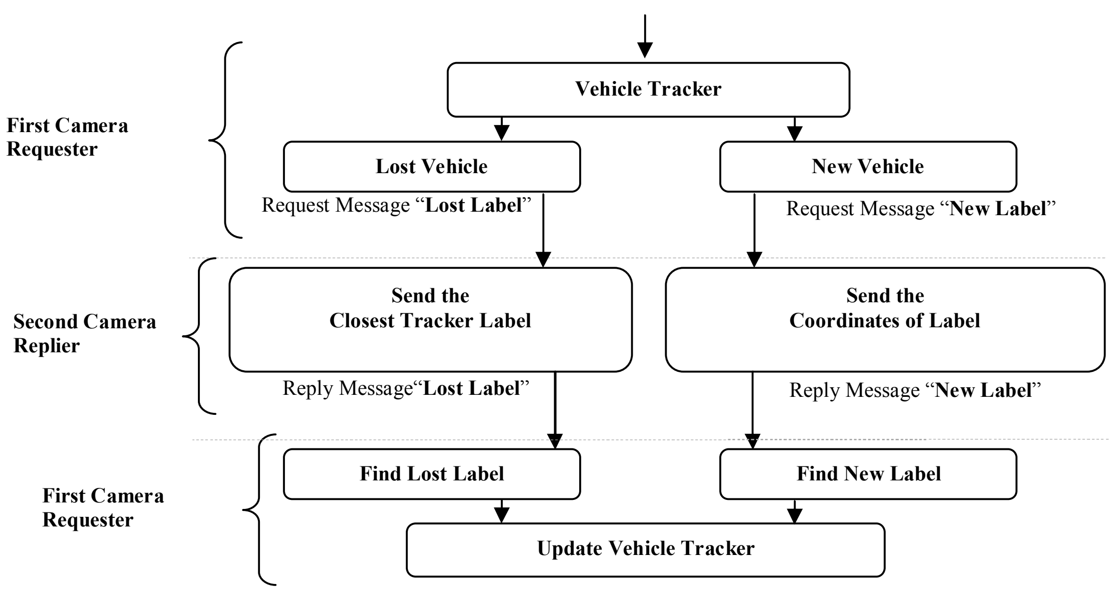

Figure5. Communication protocol between cameras

As it is shown in table 1, the format of reply messages includes the found label and the new coordinates. If an old vehicle “V_1” is not matched to any new extracted vehicles, it may have disappeared or it may be occluded by another vehicle. Thus, if V_1 disappears in the new frame, we check if there is a segment for another vehicle in another camera which has same features with V_1’s segment. If the answer is positive, vehicle is marked as occluded vehicle.

Table 1.Messages format in communication protocol

Messages Format			
New Label Request Message	Packet No.	Current Coordinates (x,y)	Temporary Label
Lost Label Request Message	Packet No.	Current Coordinates (x,y)	Lost Label
New Label Reply Message	Packet No.	New Coordinates (x,y)	Found Label
Lost Label Reply Message	Packet No.	New Coordinates (x,y)	Found Label

Another case is occurred when a new vehicle V_n is not matched to any tracked vehicles, it may be a new vehicle entering the scene, the result of improper segmentation, or a vehicle that was previously occluded and is no longer occluded. If a new vehicle V_n is detected, check whether there is a previous vehicle, whose segment overlaps with this new vehicle in both cameras. If its occlusion bit is set, then split V_1. Now, there are two vehicles V_1 and V_2 in previous frames, which were occluded now. Then,the corresponding trajectories are updated and occlusion bits reset. If V_n was not previously occluded, this is the result of improper segmentation, and should be merged. Otherwise, V_n is a new vehicle that just entered the traffic scene.

Experimental Results

In our proposed method, calibration is very important for which the transfer of relevant data between cameras is essential. When a vehicle is lost in one camera view, its coordinates is updated from other camera view and this necessitates video streams to be temporally synchronized for relevant data transfer. In experiments, we measured the accuracy of the data transfer and data updates. This measurement is computed based on the number of correct updates over total requests. The number of times a new label request or lost label request is replied correctly and corresponding tracker is correctly updated present the number of correct updates. The system accuracy for different synchronization rates is shown in Table 2.

Table2. Tracking accuracy for new and lost vehicles requests

		Tracking Accuracy					
		60s		180s		300s	
		New	Lost	New	Lost	New	Lost
Video 1	Request Messages	15	7	43	16	92	34
	Tracker updated correctly	15	6	41	14	88	29
	Accuracy	100%	85.71%	95.34%	87.50%	95.65%	85.31%
Video 2	Request Messages	25	12	74	29	112	53
	Tracker updated correctly	24	12	71	27	109	50
	Accuracy	96.00%	100%	95.94%	93.10%	97.32%	94.33%

When considering extracted trajectories in global view, we evaluate the clustering efficiency to detect the abnormal trajectories, in this case “T_p^t, True Positive Trajectories”, are those that are correctly labeled abnormal, “F_p^t, False Positive Trajectories”, are those that are incorrectly labeled abnormal and “F_n^t, False Negative Trajectories”, are those that are incorrectly labeled normal. Vehicles detection and segmentation in double view and overlapping field is shown in figure 6.Thus we need to see how the clustering has been performed in detecting the abnormal trajectories in the sequences. In table 3, the true positive and false positive values with considering the occluded trajectories in abovementioned sequences have been presented separately. Noticeably, 31 occluded trajectories are in S_1 and 23 occluded trajectories are in S_2.

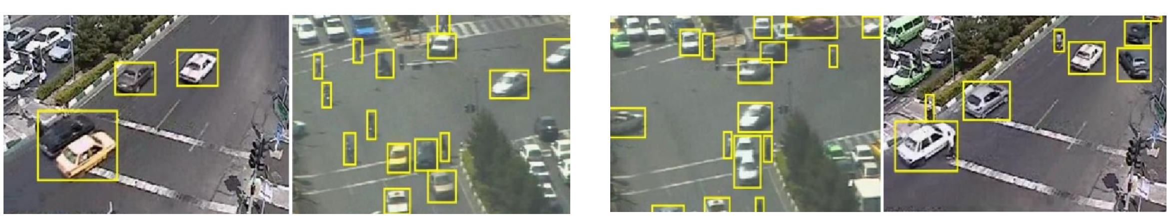

Figure6.Vehicles detection and segmentation in double view and overlapping field

Table3. True Positive and False Positive results

True Positive & False Positive					
S_1 Results (number of occlusion: 31)			S_2 Results (number of occlusion: 23)		
t seconds	True Positive	False Positive	t seconds	True Positive	False Positive
120s	8	0	120s	4	0
180s	10	0	180s	12	0
240s	17	0	240s	17	0
300s	22	1	300s	27	1
360s	33	1	360s	31	1
420s	45	2	420s	38	1
480s	52	3	480s	41	2

CONCLUSION

This paper presented a distributed multi-camera system for multiple vehicles tracking in urban traffic surveillance. In this cooperative surveillance system, two different processing units are used to process each camera and to communicate with each other directly eliminating the need for a central server and to provide scalability in computation and removing a single point of failure. Instead of transferring the extracted features to a central server, each camera performs its own tracking, keeping its own trajectories for each vehicle in the partially overlapping fields of view.A novel communication protocol was designed to coordinate multiple tracking components across the distributed surveillance system. We tested the performance on real world datasets and found that the precision and the recall results are promising. Finally, with computing tracking accuracy for different requests, New and lost label messages, we achieved at least 95.4% accuracy for new label request and 85.31% accuracy for lost label request.

ACKNOWLEGMENT

Authors are grateful to the Islamic Azad University, Science and research Branch for support to carry out this work.

REFERENCES

Babaei P, Fathy M. 2010. Multi-Camera Systems Evaluation in Urban Traffic Surveillance versus Traditional Single-Camera Systems for Vehicles Tracking. 3rd Int. Conf. on Comp. and Elect. Eng. (ICCEE 2010), vol.4, pp.438-442.

Babaei P, Fathy M. 2011. A Distributed Cooperative System forUnsupervised Vehicle's Tracking and Recognition in Urban Traffic Surveillance. 3th Int. Cof. on Machine Learning and Computing (ICMLC 2011), vol.4, pp.287-291,.

Barrois B, Hristova S, Woehler C, Kummert F, Hermes C. 2009. 3D pose estimation of vehicles using a stereo camera. In IEEE, Intelligent Vehicles Symposium.

Ben Shitrit H, Berclaz J, Fleuret F, Fua P. 2011. Tracking Multiple Objects under Global Appearance Constraints. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 137-144.

Berclaz Fleuret F, Fua P. 2009. Multiple Object Tracking using Flow Linear Programming. In Proceedings of the 12th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (Winter-PETS), pages 1-8.

Butt A, Collins T. 2013. Multi-target Tracking by Larangian Relaxation to Min-Cost NetworkFlow.Computer Vision and Pattern Recognition (CVPR), pp.1-8.

Cai Q, Aggarwal JK. 1999. Tracking human motion in structured environments using a distributed camera system. IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), vol.21.no.11, pp.1241-1247.

Chang TH, Gong S. 2001. Tracking multiple people with a multi-camera system. Proceedings of IEEE Workshop on Multi-Object Tracking, pp. 19-26.

Kelly PH, Katkere A, Kuramura DY, Moezzi S, Chatterjee S, Jain R.1995.Architecture for multiple perspective interactive video Proceedings of ACM Conference on Multimedia, pp. 201-212.

Khan S, Shah M. 2003. Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), pp.1355-1360.

Lee L, Romano R, Stein G. 2000. Monitoring activities from multiple video streams: Establishing a common coordinate frame. IEEE Transactions on Pattern Analysis and Machine Intelligence,pp. 758-768.

Leibe B, Cornelis N, Cornelis K, Van L. 2007. Dynamic 3D scene analysis from a moving vehicle. In Computer Vision and Pattern Recognition, CVPR. pp. 1-8.

Liu J, Collins T, Liu Y. 2013. Robust Autocalibration for Surveillance Camera Network. Applications of Computer Vision (WACV), pp.1-8.

Mathew TV, Krishna Rao KV. 2007.Introduction to Transportation Engineering.Chapter 39,Traffic Intersections,NPTEL.

Utsumi A, Mori H, Ohya J, Yachida M. 2000. Multiple-camera-based human tracking using non-synchronous observations. Proceedings of Asian Conference on Computer Vision, pp. 1034-1039.