Engineering Applications of Artificial Intelligence 45 (2015) 220–233

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Split-and-match: A Bayesian framework for vehicle re-identification in road tunnels

Andrés Frías-Velázquez*, Peter Van Hese, Aleksandra Pižurica, Wilfried Philips

Department of Telecommunications and Information Processing (TELIN-IPI-iMinds), Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 4 February 2015
Received in revised form
28 May 2015
Accepted 27 June 2015
Available online 25 July 2015

Keywords:
Vehicle matching
Multicamera tracking
Tunnel surveillance
Non-overlapping cameras
Trace transform

ABSTRACT

Vehicle re-identification is key to keep track of vehicles monitored by a multicamera network with non-overlapping views. In this paper, we propose a probabilistic framework based on a two-step strategy that re-identifies vehicles in road tunnels. The first step consists of splitting the re-identification problem by connecting groups of vehicles observed in different cameras using certain motion and appearance criteria. In the second step, we build a Bayesian model that finds the optimal assignment between vehicles of connected groups. Descriptors like trace transform signatures, lane change, and motion discrepancies are used to derive our probabilistic framework. Experimental tests reveal that connected groups derived from the first step are composed of 4 vehicles on average. This allow us to constrain the number of candidate matches and increase the chances of getting the correct match. In the second step, our Bayesian model succeeds in matching vehicles among candidates with very similar appearance and under uneven illumination conditions. In general, our system reports a re-identification accuracy of 92% using a nearest-neighbor matcher, and 98% using a one-to-one matcher. These results outperform previous works and encourage us to further develop our solution for other re-identification applications.

1. Introduction

Tunnel surveillance is of utmost importance for transportation authorities because the risk of being killed in a tunnel accident is twice as high as in open roads (Naussbaumer, 2007). One factor that increases the fatalities in tunnels is the risk of crashing into the tunnel walls. Another factor is the risk of fire because the chances of death by suffocation and burning increases in confined spaces. To improve the security in tunnels, multicamera surveillance systems have been extensively deployed to monitor traffic activity. The intention is to provide valuable information to security and emergency corps on hazardous situations in a timely fashion.

A downside of a multicamera surveillance system is the considerable amount of video data generated and its storage. As a tunnel length may reach up to 25 km, hundreds of cameras and human operators are needed to have a full coverage of the tunnel. To minimize the cost of the multicamera network, its maintenance, and data flow, the camera array is placed with non-overlapped Fields of View (FoV), as shown in Fig. 1. In exchange to these benefits, the identity of the vehicles is lost while passing through the blind zone due to the lack of tracking information. Consequently, it is necessary to perform an *identity handover* or re-identification to associate the tracking information from the corresponding vehicles. In practice, this re-identification is key to detect incidents in areas not covered by cameras. For example, collisions and broken-down vehicles can be detected when a vehicle is seen by one camera, but not by the next one after a certain time. Re-identification is also crucial to keep track of vehicles that pose a high risk to the tunnel safety, such as trucks carrying dangerous goods. By knowing the position of dangerous vehicles in the tunnel, proper measures can be taken in case of accident.

Even though vehicle matching has been extensively studied in the last decade, dealing with strong illumination and appearance variations remain major challenges. In road tunnels for example, illumination is not isotropically distributed, causing shades and drastic changes on the appearance of the vehicles. Moreover, head and rear lights are normally turned on, which introduce local illumination changes in the scene. All these disturbances turn the vehicle re-identification into a very challenging problem, as illustrated in Fig. 2.

In this work, we propose a Bayesian framework that jointly exploits the appearance and motion information of vehicles to perform a fast and robust re-identification. A simple example illustrating our re-identification approach is presented in Fig. 3. The proposed framework is basically composed of two steps: the first step, called *Matching Problem Decomposition* (MPD), splits the re-identification problem into several matching subproblems

* Corresponding author. Tel.: +32 92644226; fax: +32 92644295.
E-mail address: Andres.FriasVelazquez@telin.ugent.be (A. Frías-Velázquez).

http://dx.doi.org/10.1016/j.engappai.2015.06.024
0952-1976/

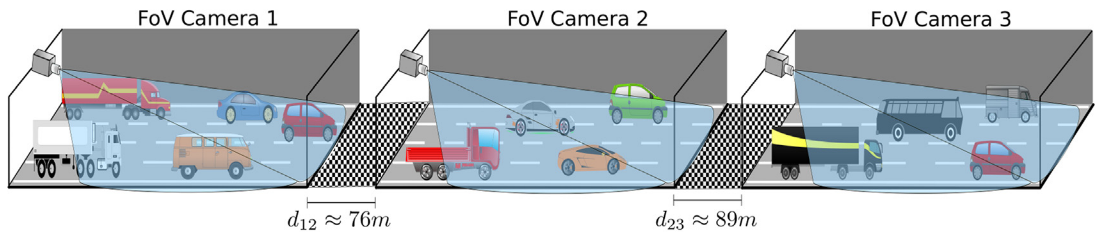

Fig. 1. Example of a multicamera setup for tunnel surveillance with non-overlapping views. Vehicle re-identification is needed to correctly associate each of the vehicles across different cameras installed in the tunnel.

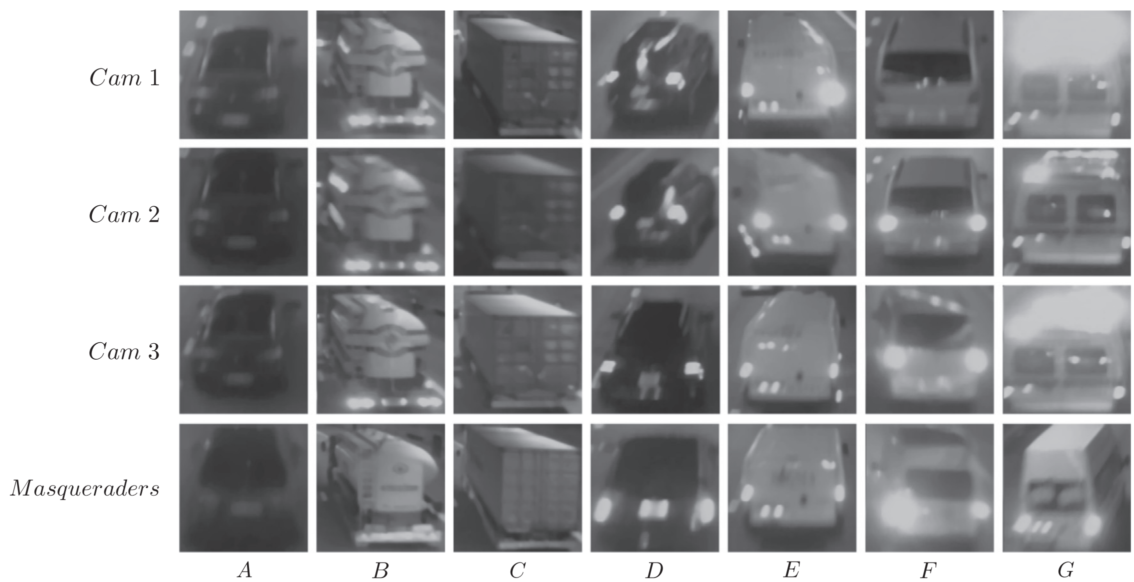

Fig. 2. Illustration of the re-identification challenges faced in a tunnel. The first three rows of each column show images of the same vehicle captured in three different cameras. The last row shows vehicles that look quite similar to their counterparts found in every column, but they come from a different vehicle. We name these distractors *masqueraders*. In columns *A*, *B*, and *C* we can see the similarity between corresponding images and masqueraders, which may lead to false alarm matches. Moreover, in columns *D* and *E* we can see that scale and pose changes may also cause mismatches. Finally, in columns *F* and *G* we can observe the impact of vehicle lights on the appearance of the vehicles.

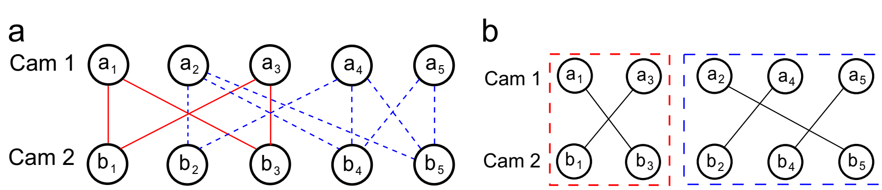

Fig. 3. A simple example of our two-step strategy performing the vehicle re-identification between adjacent cameras. In Fig. 3(a) we can see a graph that shows the first step of our strategy, called matching problem decomposition. The vertices labeled as *a* and *b* represent the vehicles observed in camera 1 and 2, respectively. The edges of the graph represent the candidate matches resulting from a statistical hypothesis test. With this information, we find the connected components of the graph like those formed with solid and dashed edges that we call matching subproblems. In Fig. 3(b) we can see another graph that represents the second step of our strategy, referred to as vehicle assignment. In this stage, a MAP estimation is used to get the best bijective assignment for each matching subproblem.

based on a statistical hypothesis test. Each subproblem represents the assignment problem of a group of vehicles observed in different cameras. The second step, called *Vehicle Assignment* (VA), finds the best bijective assignment for each matching subproblem using a Maximum-a-Posteriori (MAP) estimation. The posterior probability is built with descriptors based on appearance and motion cues.

Considering the challenging illumination conditions in tunnels, the fusion of appearance and motion information plays a key role in obtaining an optimal re-identification performance. The proposed spatio-temporal model is described in terms of several motion features related to the entry and exit points of the fields of view like lane change, speed, displacement, and transition time. Moreover, we assume that vehicles move from camera to camera following a linear acceleration model. This assumption holds for the traffic dynamics encountered in road tunnels, since the flow is normally unidirectional and the road between cameras is usually straight. To evaluate the consistency of the motion data with the kinematic model, we propose a descriptor called *spatial discrepancy*. With this descriptor we expect to discard unlikely matches and reduce the number of possible assignments.

Our appearance model is based on trace transform signatures (Petrou and Kadyrov, 2004). These image descriptors robustly deal with illumination and affine pose variations. The signatures, also called *circus functions*, are compared between images to yield a similarity measure. Also, we propose an appearance descriptor