ACIT'2017
The International Arab Conference on Information Technology
Yassmine Hammamet, Tunisia
December 22-24, 2017

Vehicle Re-identification in Camera Networks: A Review and New Perspectives

Sami GAZZAH
LATIS- Laboratory of Advanced Technology and Intelligent Systems,
ENISo, Sousse University - Tunisia
sami.gazzah@gmail.com

Najoua Essoukri Ben Amara
LATIS- Laboratory of Advanced Technology and Intelligent Systems,
ENISo, Sousse University - Tunisia
najoua.benamara@eniso.rnu.tn

***Abstract*— Vehicle re-identification aims to establish a match of the same vehicle in traffic-scenes over different cameras. Moving vehicle often changes appearance when located at different physical locations or regions duo to the lighting and angle of view variations. This paper provides a survey of the recent progress in the literature of vehicle re-identification in camera networks. The survey highlights the most recent approaches used to establish the correspondence between different video sequences provided by different cameras. The vehicles re- identification stills not a well investigated issue compared to the person re-identification. Several key issues still uncovered in the literature are well-covered. The paper points to several promoting directions for future research.**

Keywords*— Vehicle re-identification; field-of-view; Vehicle detection; styling; insert (key words)

I. INTRODUCTION

Vehicle re-identification aims first to re-identify vehicles after they leave the field-of-view (FOV) in the first camera and appear again in the FOV of another camera, and then to make correspondences across multiple camera views [1] (fig.1). Vehicle re-identification attracted increasing interest since it has various practical applications like: (1) in intelligent Transportation System (ITS) it allows to evaluate the performance of dynamic traffic systems by estimating the circulation flow and travel time. On the other hand, it provides current available information about the origin-destination matrix which is essential in transportation planning, highway toll charge model extraction [2]. (2) In security and criminal investigation processes, it allows to automatically search and track a suspicious vehicles, or road access restriction management.

Many works were led in the field of person re-identification. However, to the best of our knowledge, only few investigations addressed the problem of vehicle re-identification. Both problems can use similar processing techniques. The majorities of the re-identification processes include three key steps which are object detection in camera, object tracking frame by frame in camera, Object retrieval across cameras.

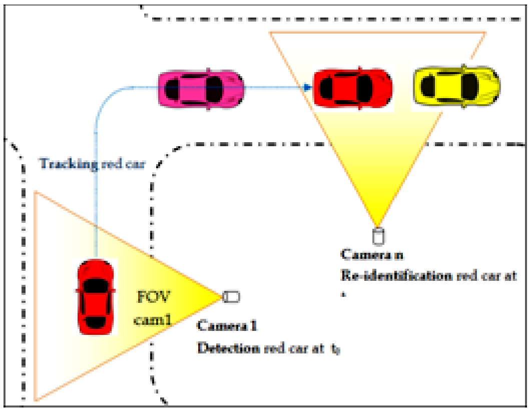


Fig. 1. Vehicle re-identification process

To track the presence of a vehicle from one frame to another in the same camera or across different views in cameras' network in wide-area surveillance and visual tracking is very challenging due to the following problems:

- Changing illumination in which vehicles are captured at different time and location. Therefore, the appearance of a vehicle may change significantly due to the illumination changes. Consequently, the color of the vehicle also would change.
- Resolutions variation. A gallery of same vehicle bonding boxes provided by one or different cameras due to the large distance between camera and vehicle subjects
- Shape of vehicle variation (body, front and tail)
- Partial or full target occlusion
- Pose changes
- Shadows and the specular reflection of car body
- Similar appearance of two different vehicles belonging to the same model.

- Intra-class and inter-class difference in appearance of vehicle in the same or different cameras.
- High volumes of activity data, different weather conditions,
- Crowded scenes, partial occlusions,
- Lighting effects such as shadows and reflections

Contributions

The spectrum of research required to achieve Vehicle re-identification at the scale envisioned above requires significant contribution along many directions. In this study, we will limit our review to the video image processing based system. The main contribution of this literature review is twofold:

— All the key aspects in the vehicle re-identification, including different approaches developed in the few existing recent works, different available databases considered and how evaluate a vehicle re-identification system, will be reviewed and discussed in order to give a clear outlook to understand a topic.

—We provide some potential directions and respective discussions about vehicle re-identification, which are open research issues.

In this paper, we first present literature survey for vehicles re-identification in Section 2. Then section 3 is dedicated to discussing some new trends in this domain. In Section 4, we describe metric evaluation and some public databases. Conclusion and future directions can be found in Section 5.

II. LITERATURE SURVEY

A. Vehicle detection trends

Due to the importance of vehicle re-identification, several approaches have been proposed. These proposed approaches basically differ from each other by the technologies involved and the sensor development. The Sensor provides attributes such as color, size, length, width, height, speed, direction of travel, date/time, and location. Accordingly vehicle re-identification systems [3] can be classified into two categories:

The first category provides continuous information, to control center, about the vehicle identity and its location in terms of space and time (GPS, GPS/GSM/GPRS). The GPS system was used firstly in shipping navigation and therefore in many other fields around the world such vehicle anti-theft tracking systems, fleet management systems and intelligent transportation systems. The GPS/GSM/GPRS system uses the GPS to get the geographic coordinates of the vehicle and then the GSM/GPRS module transmits and updates the vehicle location into a database.

The second category provides information about the vehicle identity and its location when it is located in the field-of view of one detection terminal. Inductive loops, RFID and cameras are widely used. Inductive loops consists on embedding a metallic loops in the road usually below the asphalt. Signature based system provided by detectors such as inductive loop detector and radar detector to track a vehicle across multiple detection stations. Recently, the magnetometers technology has been emerged and it replaced inductive loops. Magnetometers detect the passing vehicles when they detected a variation of the magnetic field comparing to the Earth's magnetic field [4].

Video image processing based system uses camera networks in order to provide useful images to re-identify vehicles across multiple camera views.

In the re-identification approach based RFID technology, passing vehicles are identified when an in-vehicle Tag is detected by readers installed above.

The use of the GPS and RFID technologies for re-identification task have two drawbacks. In fact, the GPS methods are accurate, but they are not well adapted to security and criminal investigation processes because they need the cooperation of the vehicle user to install hardware in-vehicle. For RFID, also specific readers must be installed above the vehicles and these readers consume many resources and they take much time. On the other hand, the installation of the inductive loops and magnetometers technologies is complicated since they need to drill a hole in the carriageway to embed it below the asphalt and also the maintenance tack is complicated and they are not accurate enough (table 1).

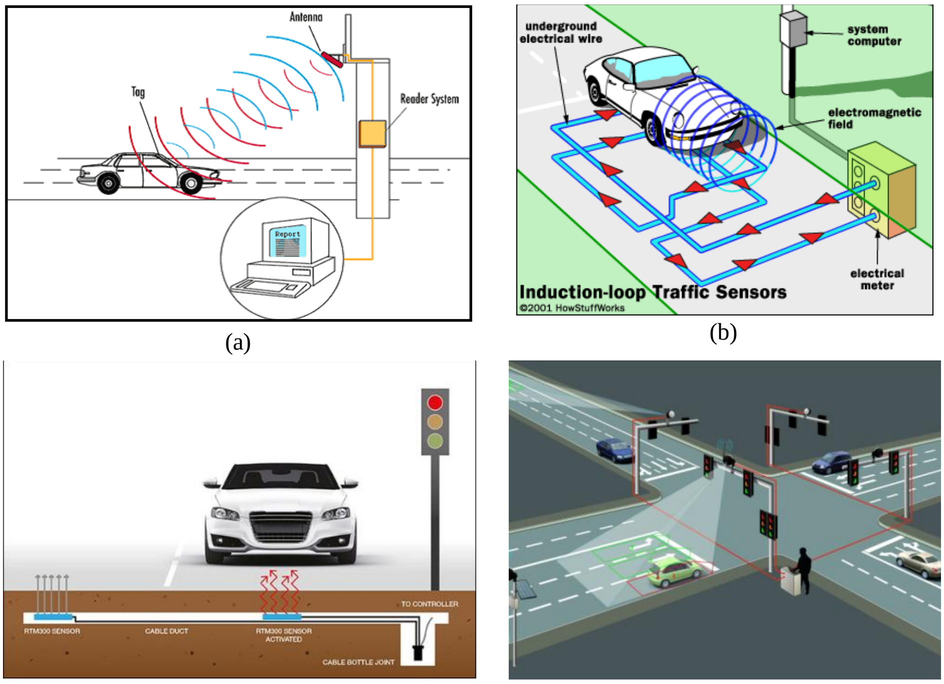
Vehicle re-identification task based on video image processing seems to be more suitable since it does not need the cooperation of the vehicle owner and usually without additional charges for installation since the cameras are usually installed everywhere around us and networked. Consequently, increasing demand to install video surveillance applications in order to take full advantage of this hardware and automatically multi-object detection, tracking and re-identification. Re-identification systems based cameras are able to process a large amount of videos automatically and rapidly with providing a promising performance, as reported in table 1.

In recent years, many terrorist attacks are held in different public places and terrorist frequently used vehicle to reach or to leave the scene of crime. A growing need for intelligent video-surveillance system able can automatic track suspect vehicles across the different streets and public places. In this context, cameras placed in different places must be networked to guarantee the exchange of detection of suspicious vehicles has emerged in different public places [5].

According to the best of our knowledge, there have been any comprehensive literature reviews on the topic of Vehicle Re-identification using Camera Networks. However, there have been a lot of reviews related to human re-identification. In recent years, a lot of works are interested in person re-identification [6], vehicle Re-Id stills on its embryonic stage since a few works are available. The use of wireless sensor networks such as camera and magnetic sensor networks allow an easy information exchange about query vehicle identification across different locations.

TABLE 1. VEHICLE DETECTION TRENDS UPON DIFFERENT TECHNOLOGIES INVOLVED [7]

	Technology	**Anonymous Data**	**Recognition rate**	
			%	**Localization precision**
Vehicle tracking	GPS	Non	100% of Vehicles Equipped with	Some meters
	GPS/GSM/GPRS	Non	100% of Vehicles Equipped with	
Vehicle Re-identification	RFID	Non	100% of Vehicles Equipped with	<1m
	Inductive Loops	Yes	30-68%	
	Camera	Non	74-99%	



(a) (b) (c) (d)

Fig. 2 : Usual technologies for vehicle re-identification

Image sources (a)http://lfavio2.blogspot.com/ (c) http://radixtraffic.co.uk (d) http://www.aldridgetrafficcontrollers.com.au

B. Vehicle signatures based sensor

Vehicle signature can be provided by the use of magnetic sensors or induction loop detection. Magnetic sensor nodes are more commonly used, among others techniques. Since the induction loop detection has low sensitivity and poor noise immunity, it is not adapted to the re-identification task [2]. In the re-identification process, a vehicle signature is used to identify a vehicle in different location by matching vehicle signatures provided by different sensors nodes. However, environmental disturbances and vehicle speed variation present a challenge for those approaches. Yin et al. [2] proposed vehicle re-identification approach based on multi-sensor spatio-temporal correlation algorithm. This algorithm is broken up into four parts: data collection and interpolation, data error estimation, waveform matching and waveform fusion. Vehicle signatures are collected using the variation of the magnetic field in different sensor nodes and transmitted to an access point via wireless sensor networks. Identification is improved by fusing information provided by multiple nodes.

Ground truth data is computed using calibrated video sequences.

C. Vehicle features based camera

The earliest vehicle re-identification systems, appears in the 80th. They used an OCR Optical Character Recognition to read the license plates captured on camera. However, in terrorist attacks and smuggling context these plates may be removed or eradicated. In the state of the art, to accomplish vehicle re-identification, significant features were intensively extracted and used. A Feature can be manually or automatically extracted

D. Hand craft features

Arth et al.[8] proposed an implementation in two networked embedded smart cameras for vehicle re-identification based on PCA-SIFT features and a vocabulary tree. Vocabulary tree represent the vehicle signatures, which are extracted from the first camera and communicated to the second ones. They achieved about 87% of successful vehicle tracking. However, the use of PCA-SIFT features as appearance-based approaches cannot correctly identify a vehicle in traffic scene due to the resemblance of vehicles and various environment factors.

In [9], the authors focus only on the vehicle re-identification step. They assume that the vehicle detection is already done and also a full-fledged 3D bonding box is previously extracted. From each 3D bounding box, they considered only the side and the front faces of a vehicle which are cropped and combined together to yield a single 2D combined images. Then, color histograms and histograms of oriented gradients (HOG) are computed from previously yielded image. Color histograms are computed using 16 bins per channel in RGB color space. A regression by linear SVM classifier was used in the re-identification task.

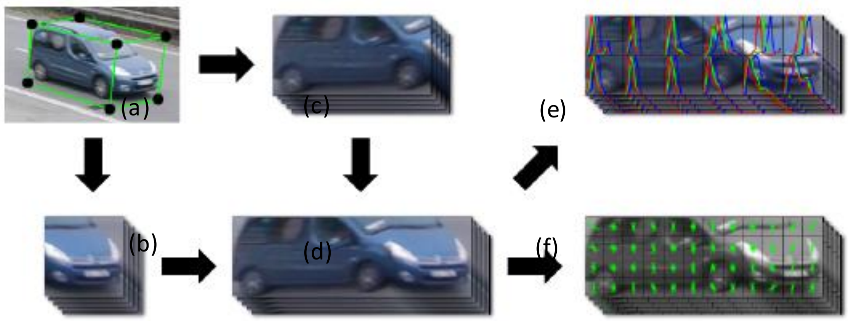


Fig. 3. (e) Color histograms and (f) HOG features extraction from (d) side and front faces of a (a) 3D bounding box vehicle (From Zapletal et al. [9])

Cormier et al. [10] presented an approach for low resolution vehicle re-identification in aerial video-surveillance system. He used a Wide Area Motion Imagery database acquired using six cameras mounted on moving airborne at 1.2 Hz low frame-rate [11] to yield monochromatic images. The problem is considered as a one-to-many re-identification problem in which a query vehicle has to be matched in a list of potential candidate. They proposed to describe vehicles using a combination of Local Binary Patterns (LBP) and Local Variance Measure (VAR) histograms computed in local grid cells. To match similarity, they proposed the uses of Hellinger distance.

E. Deep-learning networks

A new area of machine learning research has emerged, commonly called deep learning (DL), It performs extraction of key characteristics automatically without human intervention [12]. The wide use of DL architectures is driven by the recent advances in the architecture of new graphics cards and huge databases [13].

For instance, Deep-learning Neural Networks known as Convolutional Neural Network (CNN) are commonly made by multiple layers of non-linear operations compared to classical neural networks' architectures. Adding more hidden layers allows increasing the abstraction of features' hierarchy starting from low-level features extracted directly from images and fed into the input layer. These architectures are able to extract hierarchical representations from a big number of images according to their similarities. However, the training time of the CNN has been too long with the recent advances in object localizations lead to propose an improved CNN version called Region-based CNN (R-CNN) [13]. However, the uses of deep learning algorithms need a huge amount of data to automatically extract representation for each object.

As an example in [14], the authors propose an interesting database containing about 50,000 images of 776 vehicles using 20 cameras views for vehicle re-identification. Furthermore they proposed a benchmark for vehicle re-identification that cooperate three types of characteristics uses (1) Color in which authors applies bag of words (BOW) with Color name features to characterize vehicle, (2) Texture based features using Scale Invariant Features Transform (SIFT) and BOW model trained by hierarchical k-means to quantize the features and finely (3) high level semantic features extracted by deep neural networks, two models were explored AlexNet and GoogleLeNet.

Liu [14]: Use one image of a vehicle detected in one camera and looks for tracks of the same vehicle as query object in the other camera. From the query we extract the license plate image and space and time information

TABLE 2. SUMMARY OF SOME PROPOSED APPROACHES IN THE STATE OF ART

References	Re-id	Performances
[15]	CNN HOG SURF	90% to 96 42% to 45% 45% to 62%
[14]	BoW-SIFT BoW-CN	mAP= 1.85 mAP=13.95
	AlexNet GoogleLeNet	mAP=9.69 mAP=13.92
	AlexNet+BoW-CN BoW-SIFT, BoW-CN, GoogleLeNet	mAP=17.88 mAP=19.921
[9]	RGB & HoG and regression by linear SVM	ROC curves
[16]	Bayesian framework combined motion and appearance features	overall re-identification rate= 98%
[8]	PCA-SIFT	87%

III. NEW TRENDS

In this section, we focus on three new trends proposed in the literature to face the problems of vehicle re-identification, which are 1) the uses of additional contextual information: To reinforce the establishment of the correspondence between different video sequences provided, 2) transfer learning: To resolve the problem related to vehicle re-identification in a different traffic scene and 3) data fusion: To reinforce the description of data.

Contextual information

In the re-identification task, contextual information provides a valuable understanding of the factors underlying spatial and temporal aspects of the scene. Likewise to person re-identification systems [14, 18], vehicle re-id approaches can be classified into non-contextual and contextual methods. Practically, non-contextual methods rely mainly on appearances of vehicles and measure visual vehicles intra-similarities to establish correspondences, while contextual methods exploit additional contexts to improve the accuracy of the re-identification system [19]. As an example of commonly used additional information we can use: vehicle in different views, camera parameters, camera geometry information, camera topology, time between observations, license plate number, modeling changes in illumination, vehicle speed, direction and topology of the camera network. Contextual information is used in objects association stage. It is seek to reduce the research space by the temporal and spatial interrelations between cameras [17].

IV. RE-ID ALGORITHM EVALUATION

A. Metrics evaluation.

Different metrics, widely used for person re-identification, are commonly used to evaluate performance of systems based vehicle. Those metrics suppose a prior available ground truth. We use:

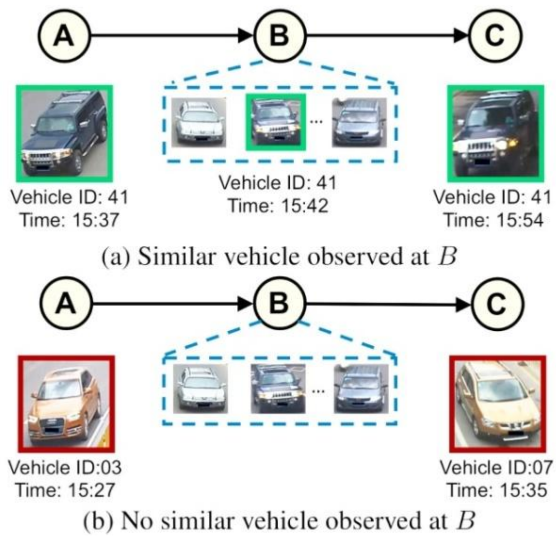


Fig. 4. Example of the uses of spatio-temporal information as contextual information [17] a vehicle with the same ID in A and C will be retained if the they observed at B and has a similar appearance and proper time.

Cumulative Matching Characteristic (CMC) curve: computes the probability that a questioned vehicle appears in different-sized candidate lists.

Moreover, we also use mean average precision: Since for each query there are N_{gt} multiple ground truths, the mean average precision (mAP) is widely used [14]. This measure considers both precision and recall to evaluate the re-identification performance.

$$AP = \frac{\sum_{k=1}^{tt} P(k) \times gt(p)}{N_{gt}} \quad (1)$$

$$mAP = \frac{\sum_{q=1}^{Q} AP(q)}{Q} \quad (2)$$

Where tt is the number of test tracks, p(k) is the precision and gt(k) is an indicator function equaling 1 if the kth result is correct and 0 otherwise.

B. Databases

Although the huge number of existing databases used for cars identification and tracking, only some ones are adapted for vehicle re-id, since we need a database in which the same vehicle must appear in more than one camera in real word traffic environment. This lack of available databases impede the advances of vehicle re-identification, according to the table 3, the majority of databases was available since 2016.

Peng et al. [15] collected vehicles datasets using six non overlapping cameras in the campus university. It contains six videos with 60 minutes duration in the resolution of 1920x1080 with 25 frames per seconds (fps).

[11] A public dataset provided by the US contains imagery from a large format monochromatic electro-optical platform. The sensor is comprised of six cameras, with each of them oriented to maximize coverage but still allow overlap. It contains about 1,537 stitched images in National Imagery Transmission Format NITF 2.1, JPEG 6.2 pyramid stitched images, raw camera data and derby index database of the NITF files and the image truth. In [10] this dataset was used for low resolution vehicle identification since each vehicle covers about 10x20 pixels (fig.5).

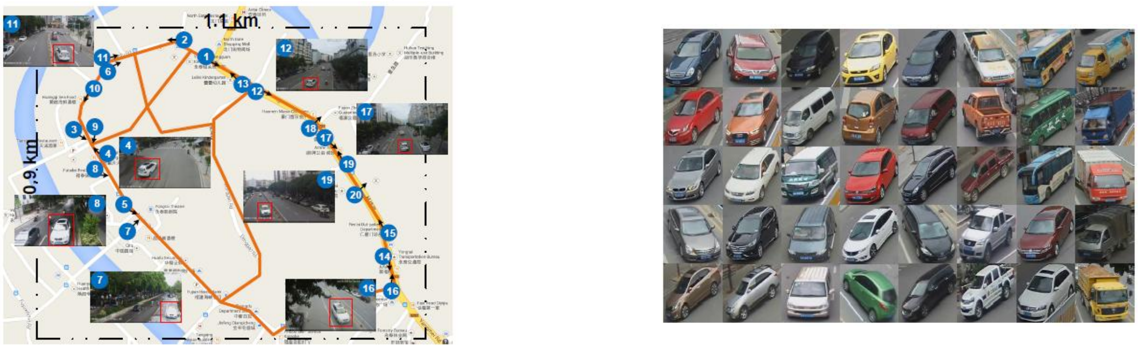
Fig. 5. Image sample from [11]

TABLE3 .AN OVERVIEW OF PUBLICLY AVAILABLE DATA SETS.

Data set	Year	#camera	Overlapping	Web link (if available)
[15]	2016	6	no	
“VeRi” [14]	2016	20	yes	https://github.com/VehicleReId/VeRidataset
[16]	2015	3	no	
[8]	2007	2	no	
[9]	2016	2	no	
[10]	2016	6*	yes	https://www.sdms.afrl.af.mil/index.php?collection=wpafb2009

*Six cameras were used to create Wide Area Motion Imagery

The VeRi dataset [14] consist of videos of vehicles in different color and type. It contains over 50,000 images of 776 vehicles a 1 km2 area in 24 hours is considered using 20 cameras views. The datasets were collected along a circular road. The scenes of the cameras include two-lane roads, four-lane roads, and crossroads. All cameras are in the resolution of 1920 x 1080 with 25 fps. Besides, there are overlaps for part of the cameras. It contains 619 vehicles, 7,021 tracks and 40,395 bounding box (BBox) images in which only moving vehicles were annotated and most vehicles have more than five tracks over the whole camera network. Cars are labeled with varied attributes such as BBox, colors, types and models. It is also labeled with license plates and spatiotemporal information. Each vehicle is captured by 2 until 18 cameras in different viewpoints, illuminations, resolutions, and occlusions, which provides high recurrence rate for vehicle Re-Id in practical surveillance environment fig. 6).



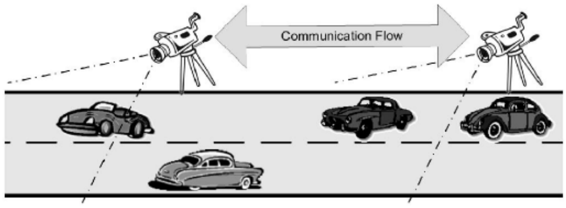
(a) (b)

Fig. 6. The VeRi dataset [14] consist of videos of vehicles in different color and type (b). The urban surveillance environment covers 1 km2 area using 20 cameras views (a).

[20] proposed three manually annotated video sequences provided by three non overlapping cameras in a tunnel. Images are in the resolution of 720x576. In the annotation process each vehicle is delimited by a bounding box and appears at least in 15 frames. Database is composed of around 460 vehicles, 300 were used for training and 160 were used for testing. Vehicles are mostly cars and trucks.

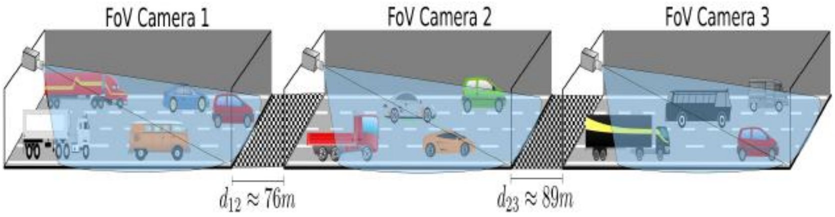
[16] collected vehicles datasets using three non overlapping cameras in a tunnel. It contains three gray-scale videos with 11 minute durations in the resolution of 720 x 576 with 25 frames per seconds (fps). It contains 564 manually annotated vehicles, each of them observed in the three cameras (fig. 8).

(a)



(b)

Fig. 7. Samples of views (a) provided by (b) two non-overlapping cameras [8].



(d)

Fig. 8. Example of views provided by three cameras installed in a tunnel (a),(b),(c) and multi-camera setup for tunnel surveillance with non-overlapping views (d). [16].

V. CONCLUSION

This paper provides a survey of the recent advances in the literature review of vehicle re-identification in urban surveillance scenario. Many approaches were proposed and differ basically from each other by the technologies involved and the sensor development. The use of wireless sensor networks such as camera and magnetic sensor networks allow an easy information exchange about query vehicle identification across different locations. This review highlights the approaches based on cameras which are more commonly used. Since person re-id is widely aborted in the state of the art, we have reviewed how to exploit for vehicle re-identification. The survey highlights the most recent approaches used to establish the correspondence between different video sequences provided by different cameras. Recently, New emerged approaches based on deep learning methods were widely used for vehicles re-identification task due to being independent of handcrafted features. Contextual information provides a valuable understanding of the factors underlying spatial and temporal aspects of the scene that allows understanding and reinforcing the establishment of the correspondence between different video sequences provided.

VI. REFERENCES

[1] Assari, S. M.; Idrees, H. & Shah, M. Human Re-identification in Crowd Videos Using Personal, Social and Environmental Constraints Computer Vision textendash ECCV 2016, Springer International Publishing, 2016, 119-136.

[2] Yin T., Hong-hui D., Li-min J. and Si-yu L. “A vehicle re-identification algorithm based on multi-sensor correlation”, Journal of Zhejiang University SCIENCE C, 2014.

[3] Shin –Ting J., Real time Re-identification System for freeway performance measurement. Doctorate thesis, University of California, 2007.

[4] Christou, C. T. & Jacyna, G. M. Vehicle detection and localization using Unattended Ground Magnetometer Sensors 2010 13th International Conference on Information Fusion, IEEE, 2010.

[5] Saghafi, M. A.; Zaman, H. B.; Saad, M. H. M. & Hussain, A. Review of person re-identification techniques IET Computer Vision, Institution of Engineering and Technology (IET), 2014, 8, 455-474.

[6] Zheng, L.; Yang, Y. & Hauptmann, A. G. Person Re-identification: Past, Present and Future 2016.

[7] PITTON A-C; Contributions à la Ré-Identification de Véhicules par Analyse de Signatures Magnétiques Tri-Axiales mesurées par une Matrice de Capteurs, PHD thesis, university of Gronoble/France, 15th january 2015.

[8] Arth, C.; Leistner, C. & Bischof, H. Object Reacquisition and Tracking in Large-Scale smart Camera Networks 2007 First ACM/IEEE International Conference on Distributed Smart Cameras, IEEE, pp. 156-163, 2007.

[9] Zapletal, D. & Herout, A. Vehicle Re-identification for Automatic Video Traffic Surveillance 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2016.

[10] Cormier, M.; Sommer, L. W. & Teutsch, M. Low resolution vehicle re-identification based on appearance features for wide area motion imagery 2016 IEEE Winter Applications of Computer Vision Workshops (WACVW), IEEE, 2016

[11] U.S. AFRL. Wright-Patterson Air Force Base (WPAFB) dataset. https://www.sdms.afrl.af.mil/index.php?collection=wpafb2009 , 2009.

[12] Y. Bengio , A. Courville, and P. Vincent"Representation Learning: A Review and New Perspectives," arXiv preprint arXiv: 1206.5538, 2014.

[13] S. Gazzah, A. Mhalla, N. Essoukri Ben Amara: Vehicle detection on a video traffic: review and new perpectives. 7th International Conference: Sciences of Electronics, Technologies of Information and Telecommunication, SETIT'17, Hammamet, Tunisia, December 18-20, 2016.

[14] Liu, X.; Liu, W.; Ma, H. & Fu, H. Large-scale vehicle re-identification in urban surveillance videos 2016 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2016.

[15] Peng, J.; Shen, T.; Wang, Y.; Zhao, T.; Zhang, J. & Fu, X. Continuous Vehicle Detection and Tracking for Non-overlapping Multi-camera Surveillance System, Proceedings of the International Conference on Internet Multimedia Computing and Service - ICIMCStextquotesingle16, ACM Press, 2016.

[16] Frias-Velázquez, A.; Hese, P. V.; Pižurica, A. & Philips, W. Split-and-match: A Bayesian framework for vehicle re-identification in road tunnels Engineering Applications of Artificial Intelligence, Elsevier BV, 2015, 45, 220-233.

[17] Cho, Y. J.; Park, J. H.; Kim, S. A.; Lee, K. & Yoon, K. J. Unified Framework for Automated Person Re-identification and Camera Network Topology Inference in Camera Networks, 2017

[18] A. Bedagkar-Gala and S. K. Shah. A survey of approaches and trends in person re-identification. Image and Vision Computing, 2014.

[19] Shen, Y.; Xiao, T.; Li, H.; Yi, S. & Wang, X. Learning Deep Neural Networks for Vehicle Re-ID with Visual-spatio-temporal Path Proposals. arXiv preprint arXiv:1708.03918v1, 2017.

[20] Rios-Cabrera, R.; Tuytelaars, T. & Gool, L. V. Efficient multi-camera vehicle detection, tracking, and identification in a tunnel surveillance application Computer Vision and Image Understanding, Elsevier BV, 2012, 116, 742-753.