International Journal of Transportation Science and Technology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

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

Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles

Giuseppe Guido *, Vincenzo Gallelli, Daniele Rogano, Alessandro Vitale

Department of Civil Engineering, University of Calabria, Via Bucci cubo 46/B, Rende 87036, Italy

ARTICLE INFO

Article history:
Received 15 September 2016
Received in revised form 23 December 2016
Accepted 28 December 2016
Available online xxxx

Keywords:
UAV
GPS
Vehicle tracking
Traffic data
Video processing

ABSTRACT

This paper presents a methodology for tracking moving vehicles that integrates Unmanned Aerial Vehicles with video processing techniques. The authors investigated the usefulness of Unmanned Aerial Vehicles to capture reliable individual vehicle data by using GPS technology as a benchmark. A video processing algorithm for vehicles trajectory acquisition is introduced. The algorithm is based on OpenCV libraries. In order to assess the accuracy of the proposed video processing algorithm an instrumented vehicle was equipped with a high precision GPS. The video capture experiments were performed in two case studies. From the field, about 24,000 positioning data were acquired for the analysis. The results of these experiments highlight the versatility of the Unmanned Aerial Vehicles technology combined with video processing technique in monitoring real traffic data.

Introduction

Vehicle tracking data can be useful for the timely and efficient control and management of traffic, and may constitute a verifiable real-world platform for comparing traffic simulation outputs. The acquisition of vehicle tracking data is both expensive and technically complex, frequently requiring the deployment of costly traffic monitoring systems. Both "Infra structure-based" and "Noninfrastructure-based" techniques are currently used to obtain traffic data worldwide.

"Infrastructure-based" techniques involve intrusive detector technologies, such as inductive loops (Sun and Ritchie, 1999; Coifman et al., 2003), pneumatic road tubes (Federal Highway Administration, 2007; McGowen and Sanderson, 2011), magnetic detectors (Kwong et al., 2009; Haoui et al., 2008), piezoelectric (Li and Yang, 2006). "Noninfrastructure-based" techniques include nonintrusive detector technologies, such as microwave radar (Zwahlen et al., 2005; Ho and Chung, 2016), passive acoustic (Nooralahiyan et al., 1998; Tyagi et al., 2012), infrared (Hinza and Stilla, 2006; Grabner et al., 2008), ultrasonic (Kim, 1998; Song et al., 2004), beacon (Sanguesa et al., 2013; Bai et al., 2013).

Generally, data obtained by detector technologies are somewhat aggregate in nature and do not provide an effective record of individual vehicle tracks in the traffic stream. This limits the use of these data in analyzing individual driving behavior and calibrating and validating simulation models.

Recently, two others "Noninfrastructure-based" techniques have been applied more and more frequently to observe traffic flow conditions: probe vehicle data acquirement and video image processing.

Peer review under responsibility of Tongji University and Tongji University Press.
* Corresponding author.
E-mail address: giuseppe.guido@unical.it (G. Guido).

http://dx.doi.org/10.1016/j.ijtst.2016.12.001
2046-0430/

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

The first allows observing the trajectories of individual vehicles yielding Lagrangian measurements. These measurements can be obtained through on-board tracking devices that provide real-time information on individual vehicles.

Their application ambits are real-time traffic operations monitoring, incident detection, and route guidance applications. Probe vehicle data collection systems can be grouped in five types:

- Automatic Vehicle Identification systems (AVI), in which vehicles are equipped with electronic tags (Smalley et al., 1996);
- Signpost-Based Automatic Vehicle Location (AVL), particularly used by transit agencies, in which probe vehicles communicate with transmitters mounted on existing signpost structures (Polk and Pietrzyk, 1995);
- Ground-Based Radio Navigation, in which data are collected through communication between probe vehicles and a radio tower infrastructure (Vaidya et al., 1996);
- Cellular Geo-location based on cellular telephone call transmissions (Sumner et al., 1994; Astarita et al., 2006);
- Global Positioning System (GPS), in which probe vehicles are equipped with GPS receivers (Choi and Chung, 2001; Yim and Cayford, 2001).

However, this technique does not allow macroscopic traffic data to be obtained, such as traffic volume and density if not coupled to stationary detectors to estimate traffic conditions.

On the other hand, video image processing provides a low cost nonintrusive procedure for capturing individual vehicle operations over time, and as such, provides a useful tool for obtaining observational data for traffic control, and calibration and validation of traffic simulation models.

Video-based vehicle tracking methods can be classified into six categories which include model-based tracking (Koller et al., 1993; Baker and Sullivan, 1992; Schlosser et al., 2003), region-based tracking (Zhang et al., 1993; Oh et al., 2009; Huang and Yen, 2004), active contour-based tracking (Koller et al., 1994), feature-based tracking (Kanhere and Birchfield, 2008), Markov random field tracking (Kamijo et al., 2001) and color and pattern-based tracking (Chachich et al., 1996; Zehang et al., 2002).

However, video capture data are subject to error, especially under conditions of significant camera angle distortion. Recently the availability and deployment of traffic drones have assisted the acquisition of accurate vehicle tracking profiles based on video inputs. This is rather new technology, however, and the level of accuracy of the resultant tracks has not been fully established under varying road and traffic conditions.

Among nonintrusive techniques, Unmanned Aerial Vehicles (UAV) technologies have recently been improved on use for tracking vehicles' trajectories and estimating traffic parameters.

UAVs are quickly gaining popularity worldwide due to their dramatic increase in commercial use; in fact, although they were developed for military purposes (they are still used for a number of missions, including reconnaissance and attack role), UAVs are growing in the enterprise sector as rules allow new use cases in the US and EU, the two biggest potential markets for enterprise UAVs.

UAV-based systems have many advantages compared to manned air vehicles, first of all the cost of use related to the low purchase, management and operation costs. They can yield high resolution images useful for traffic analysis based on video image processing, but their applications for this field are limited and influenced by some factors affecting their performance. Among them, weather conditions should be mentioned (e.g. wind, rain, snow, electric and magnetic fields), technical instrumental problems (e.g. low battery duration, battery life, limited UAV payload), physical obstacles (e.g. vegetation, buildings, urban canyons), regulatory issues (e.g. limited take-off mass, limited wing surface, limited wing loading, no-fly zones, pilot certificates).

This paper presents a methodology to extract traffic data by using Unmanned Aerial Vehicles. The proposed methodology implies the use of a combined technique involving an Unmanned Aerial Vehicles (UAV) image acquisition technology and a video processing algorithm. Moreover, this paper intends to demonstrate the usefulness of Unmanned Aerial Vehicles to acquire reliable traffic data and to provide useful information on driving behavior parameters for individual drivers. Accuracy in the traffic data evaluation is assessed by comparing individual vehicle trajectories as extracted from the video with the corresponding profiles obtained from baseline GPS referenced values for the same trajectories. The study described in this paper has two specific objectives:

- to introduce and describe a methodology for tracking vehicles based on Unmanned Aerial Vehicles (UAV) image acquisition technology;
- to assess the accuracy of this methodology in evaluating individual vehicle paths.

The paper is organized according to the following structure: Section "Literature review" provides a state of the art of the traffic data collection techniques, with emphasis on UAVs and new technologies; Section "Methodology" introduces the methodology used to extract data from traffic stream though UAVs applications; Section "Case studies and results analysis" illustrates an application of the methodology described in the previous section to case studies; the results provided by the proposed methodology are discussed in Section "Summary and conclusions"; Section 6 summarizes the major conclusions of the study.

Literature review

The technologies employed for the acquisition of traffic data, in past years, were mainly represented by dedicated equipment, such as inductive loops, magnetic detectors, piezoelectric, microwave radar and infrared, all characterized by high installation and maintenance costs. Furthermore, this equipment can only ensure vehicle data onto certain sections of the road, which is an important limit to traffic analyses.

With the advent of mobile technology and the resulting diffusion of GPS equipped smartphones in the road users' community, the approach to mobile devices as a tool for traffic data extraction has been widely used by several researchers.

Guido et al. (2014) presented an innovative approach to collect vehicle data using GPS sensors on smartphones. They demonstrated high accuracy of mobile devices in the estimation of a probe vehicle speed comparing smartphones data to benchmark values obtained using a high frequency video V-Box (Race Logic®) mounted on the same vehicle. Their study yielded GPS confidence intervals of instantaneous vehicle operating speeds, subject to varying satellites signal disruptions en-route.

Handel et al. (2014) presented a framework based on smartphones as a measurement system for road vehicle traffic monitoring. Herrera et al. (2010) proposed a GPS smartphone probes based system for traffic monitoring, using data provided by the cellular networks of many urban areas. Guido et al. (2012) introduced an empirical procedure for the estimation of indicators referring to road safety performance based on vehicle interactions in real-time using new generation smartphones with GPS capabilities.

Over the last decade, aerial image processing has been applied to the field of Intelligent Transport Systems (ITS) and become a very popular research topic mainly for the increased data availability. Aerial images provide better perspective and can cover a large area in every frame, with the advantage of being mobile, thereby creating a dynamic data collection system. Thus, in the past few years, unmanned aerial vehicles (UAVs) has been an active area of research for different purposes, such as damage assessment, 3D mapping and traffic inspection.

A very important project on the employment of UAVs in traffic problems is the Wallenberg Laboratory for Information Technology and Autonomous Systems project (Doherty et al., 2000) on the technologies and functionalities necessary for the successful deployment of a fully autonomous UAV operating over road and track networks (Granlund et al., 2000). The results of the project have promoted the development of reliable software and hardware architectures, sensory platforms and sensory interpretation techniques, efficient inferencing and algorithmic techniques to access geographic, spatial and temporal information and simulation modeling tools supporting an autonomous navigation of UAVs at different altitudes (including autonomous take-off and landing).

Puri (2005) made a survey of the research activities going on in several universities around the world in the area of application of UAVs in traffic surveillance. From this review, it emerged that UAVs can be very useful and successful for traffic surveillance owing to their maneuverability as compared to ground vehicles.

Considering the employment of UAVs in simulation, Puri et al. (2007), in a Ph.D. dissertation, proposed the use of unmanned helicopters to obtain detailed traffic information on real-time. In particular, an approach is presented to generate statistical profiles of obtained data, which provides traffic engineers valuable information about the traffic characteristics. Moreover, the authors used a traffic simulation software to analyze the effects of parameter variations on the output. They calibrated the simulation model by minimizing the difference between the simulated results and the real world data expressed in terms of measure of effectiveness (MOEs) such as volume, density, travel-time and delay. They found that simulation model performs better based on proper calibration of MOEs. Results showed that UAVs are particularly suitable to obtain vital information about the traffic pattern and their employment can be utilized to develop more accurate simulation models, which can be used by traffic engineers for taking informed decisions.

Confirming the previous research, Barmpounakis et al. (2016) demonstrates that, despite few technological obstacles, these systems can be employed for real time traffic monitoring and for the extraction of vehicular trajectories from a video image processing system to collect vehicle traffic data.

Salvo et al. (2014) proposed a method to evaluate traffic flow conditions in urban areas using videos through UAV. They evaluated UAV technical characteristics and developed a methodology to extract kinematic quantities of vehicular flow, comparing the measured data onto the Region Of Interest with the models of macrosimulation. With this paper the authors stated that UAVs data output could be used as input data for calibrating models of microsimulation, in order to have simulations that are more representative of reality. Moreover, the UAV's, according the authors, allow the classification and the counting of vehicles in transit with a reduced time and at low cost capturing the dynamics of the traffic at any point in urban areas.

In 2012, Braut et al. (2012) estimated traffic flow parameters of a road intersection through a video image processing technique using an Unmanned Aerial Vehicle hovering at the heights of about 50 m. In particular, first, they assessed the accuracy of OD matrices estimated from the video acquired by a fixed camera from a tall building, based on a straightforward background modeling technique. Secondly, they applied a video stabilization technique from a hovering airship, in order to extend the applicability of fixed-camera approaches to UAV-based applications. The proposed methodology highlights encouraging results.

Lee et al. (2015), treated the applicability of small quadcopter drones for traffic surveillance and roadway incident monitoring. In order to examine the effectiveness of quadcopter drones for traffic surveillance and incident monitoring a series of pilot tests were conducted. The authors demonstrated that compared to existing traffic data collection practices, quadcopter drones appear highly beneficial as they are capable of covering a wide range of data collection sites. Pilot tests also demonstrated how the quadcopter drone is suitable for real-time roadway incident monitoring, though up to 20 s of communication lag was observed under low cellular signal strength.

Coifman et al. (2004, 2006) investigated the employment of UAVs for a wide range of transportation operations and planning applications for a series of flights: incident response, monitor freeway conditions, coordination among network of traffic signals, traveler information, emergency vehicle guidance, track vehicle movements towards an intersection, measurement of typical roadway usage, estimate Origin–Destination (OD) flows.

Rango et al. (2006) investigated data acquisition through UAV applications, in particular on a combination of different levels of sensing data to provide more comprehensive information related to the distance between the camera and test site. They focused the investigation into the utilization of the most technologically advanced sensors for monitoring and on the minimization of costs in order to maximize simplicity for monitoring purposes.

Doğan et al. (2010) proposed a real-time vehicle speed estimation of only moving objects using side view images taken with a video camera. The procedure used in their research involves two main steps. During the first step, they select a certain number of points of the vehicle to track them on at least two frames later. The second step consists in using the displacement vectors of the tracked points to compute speed values.

Recent studies on vehicle data extraction through computer vision technology have been developed.

Lai et al. (2001) presented a method to extract moving vehicles from traffic image sequences using a set of coordination mapping functions. In their study, they estimate vehicle lengths of within 10% in each instance, deriving coordination mapping functions of a calibrated camera model.

Rad and Jamzad (2005) performed a similar work developing an application able to count and classify vehicles and to identify the lane-changing events by tracking them. In their study, the background subtraction method combined with morphological operations has been applied in order to identify moving vehicles in regions.

Apeltauer et al. (2015) proposed a methodology to extract and tracking moving vehicles from aerial video data acquired with UAV. Their approach shows very good results in the automatic extraction of vehicles' trajectories for several traffic analyses.

In Cheng et al. (2012), they performed a pixel-wise detector system for vehicles identification, employing dynamic Bayesian network for the classification step.

In Papageorgiou et al. (1998), an alternative approach based on a feature set based on Haar wavelets (Haar, 1910) instead of the usual image intensities was discussed. Viola and Jones (2001) developed the so-called Haar-like features by adapting the using of Haar wavelets. They proposed a machine learning approach to object detection characterized by a high detection rate for a very short amount of time. The learning process is based on AdaBoost, a boosting algorithm that reduces extremely the computational time.

The methodology was extended to a study made by Lienhart and Maydt (2002), in which the authors introduced the diagonal features method. They increased the dimension of the set of features in a successfully attempt to increase the accuracy in object detection introducing the concept of a tilted Haar-like features.

Methodology

In order to assess the validity of Unmanned Aerial Vehicles to acquire reliable traffic data and vehicle paths, an instrumented vehicle was equipped with 20 Hz GPS receiver. GPS speed and position data of the instrumented vehicle were used as a benchmark values to assess the accuracy of data determined using video image processing technique. The biggest challenge of the experiment was to automatically identify the instrumented vehicle paths of the traffic stream among all moving vehicles and objects. The problem was solved applying the Haar classifier approach. The details of the methodological approach adopted in this paper are described in the following subsections.

The employed technology

A FIAT Doblò was used as instrumented probe vehicle; a target useful for uniquely identifying the vehicle was placed on the roof barycentrically with respect to the size of the vehicle (Fig. 1).

Equipment used to track this vehicle and to assess its speed and position includes a UAV drone with eight propellers and a video camera (Fig. 1), and a Racelogic® V-Box unit on-board (Fig. 2). The video camera is able to capture videos up to 4 k, at a frame rate of 23 fps, and photos with a 12 megapixel resolution. The Racelogic® V-Box unit on-board incorporates a 20 Hz GPS receiver that was calibrated to yield speeds within a 0.1 km/h margin of error and 0.05% accuracy in distance (less than 50 cm per km).

Tables 1 and 2 summarize the main features of the UAV and the V-Box unit.

Fig. 1. UAV equipped with a video camera.

Video processing stage

The methodology used for data analysis is essentially characterized by three phases:

1. data collection;
2. data processing;
3. vehicles data extraction.

The first stage refers to video capturing performed by UAV, based on the defined elevation, camera axis angle, video resolution and GPS point coordinates. These data are essential during the camera calibration phase, because they are the basic parameters for initializing the video-processing procedure.

The quality of the georeferencing process of the frames acquired by UAV was improved correlating pixels in the imagery with a set of corresponding GPS coordinates. The correction was obtained by creating correlations between raw images acquired by the UAV camera and a set of well-known location GPS points using a desktop GIS as shown in Fig. 3. This method refers to Ground Control Points (GCPs). Ground Control Points, which have known coordinates, increase significantly the absolute accuracy of the analysis and allow the mapping of video pixels.

Ground Control Points must be visible in all frames and placed homogeneously in the area of the analysis, in order to correct orientation and errors in photo geometry of every frame. These errors are generally caused by UAV tilt or camera lens problems, and therefore the number of GCPs is directly proportional to the desired level of accuracy in relation to the distortion of each frame of the video.

Fig. 2. Instrumented probe vehicle.

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

Table 1
Micro-drone technical features.

Technical features	Description
Body	Carbon fibre
Rotors	Brushless motors
Payload	500 g
Power	LiPo battery
Flight time	About 8 min
Wind speed	Until to 5 m/s
Operative temperatures	From 5 °C to 40 °C
GPS	Yes
Drive mode	Manual/flight planning
Video quality	Full HD (1920 × 1080 pixel)

Table 2
V-Box technical features.

Technical features	Description
GPS data logger	20 Hz
Video input	4 cameras
Cameras' resolution	580 L and 420 L
Info	Real-time graphics (bar graphs, circuit plots, lap times)
Accuracy speed	0.1 km/h
Accuracy distance	0.05% (<50 cm per km)

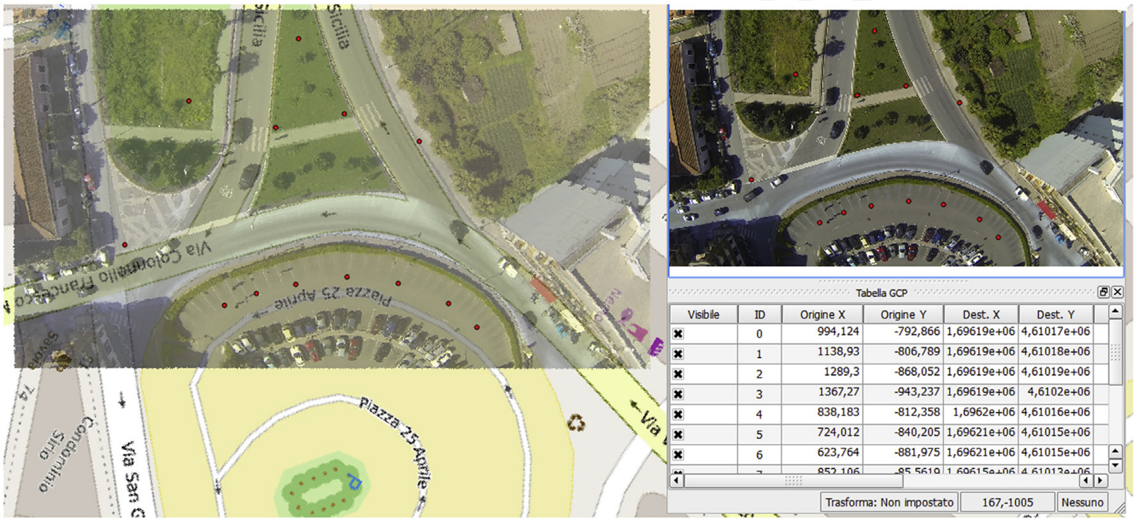

Fig. 3. Ground Control Points georeferencing process.

The process of correction consists in a regression equation that resolves the association of every pixel to real world coordinates. In the simplest case, a minimum of three GCPs generates a linear equation, which allows gaining good results in terms of low image distortion, but the photo-geometry distortion cannot be corrected. In order to reach better accuracy results, a second-order or a third-order equation (with respectively 6 or 10 GCPs required) is necessary for more complex photo-geometry distortion cases. For the experiment, at least 10 GCPs are placed in the area recorded by UAV camera.

Several issues caused by the equipment can affect the quality of the video acquired by UAV. The first effect is barrel distortion, caused by the different size of the lens field of view compared to image sensors in many cameras embedded in UAVs. Another problem that can add noise to the video is the movement caused by wind and engine vibrations. In order to maximize the effectiveness of the tracking process, some appropriate filters were applied through the OpenCV library (Bradski and Kaehler, 2008; OpenCV Library, 2015).

In particular, since the Haar classifier approach gives the best results with a fixed camera, a video stabilization filter was applied, which consists in a rigid Euclidean transformation, based on three parameters, *tx* and *ty* that define the pixels trajectory, and αxy as the angle. The correction is performed by smoothing out the identified trajectory of each pixel using a sliding average interval; the transformation is applied frame by frame (Fig. 4).

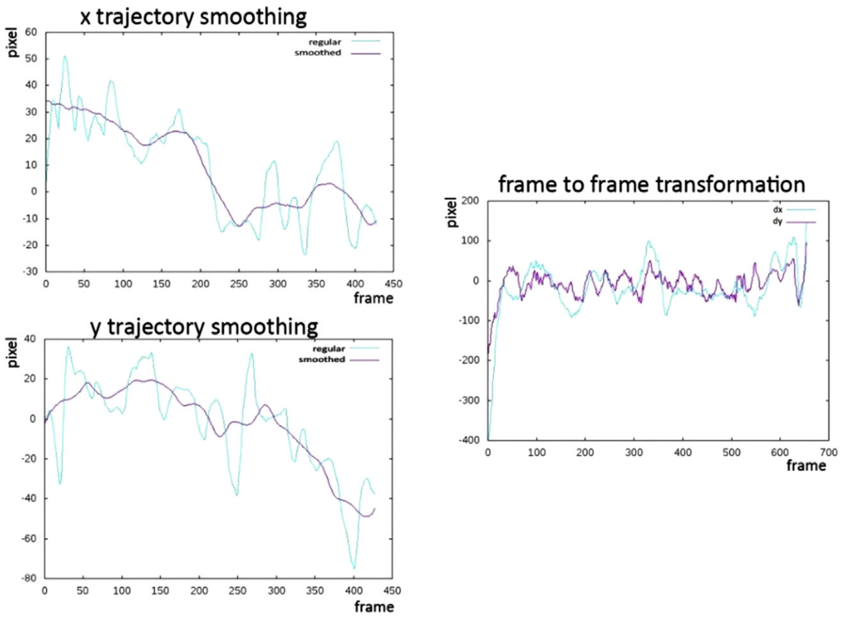

Fig. 4. Pixel correction in video frames stabilization stage.

Furthermore, a second filter was applied for the color space conversion (*cvtColor*, OpenCV), which converts each frame from one color space to another. In this study, using the command "*cv2.cvtColor (frame, cv2.COLOR_BGR2GRAY)*", all frames were converted to grayscale to reduce file size and increase the overall processing performance since color has no bearing on the vehicles detection process.

Finally, after the color correction, a Gaussian-blurring filter was applied in order to smooth all frames of the acquired video. The application of the last filter is crucial since no two frames are identical, owing to sensible variations in the camera sensors that produce different intensity values in some pixels. This filter smoothens out high frequency noise that might compromise the vehicle tracking process.

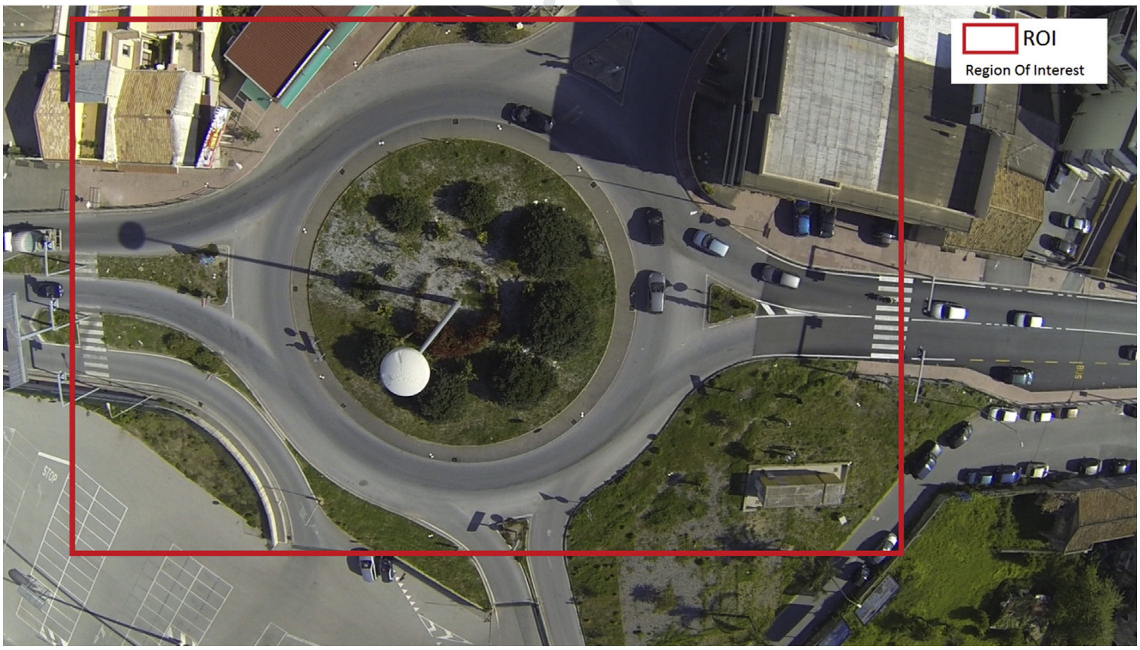

Fig. 5. Region Of Interest (ROI) definition.

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

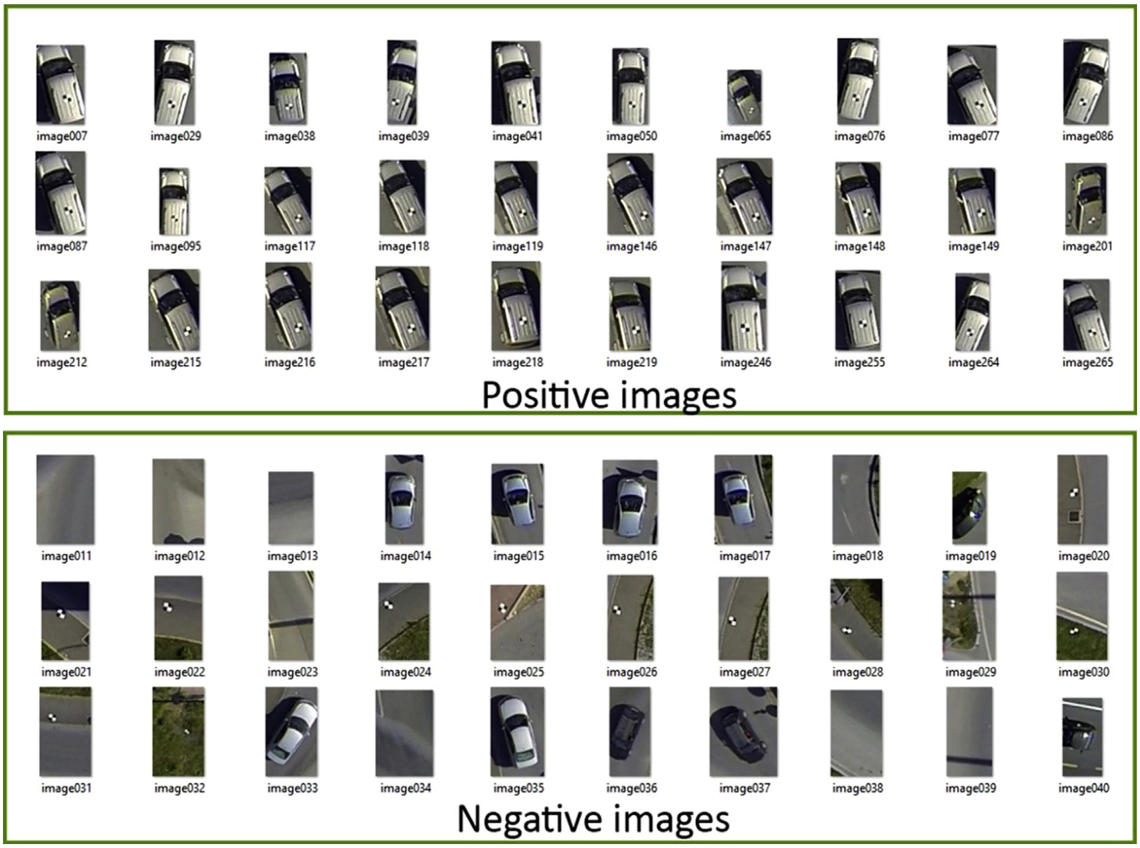

Fig. 6. Positive and negative images sample set for the training.

Once all frames acquired by the UAV camera were corrected by any distortion and each pixel was georeferenced, a Region Of Interest (ROI) was defined (Fig. 5). The Region Of Interest is a portion of the frame where the analysis was focused in order to reduce the computational load required for the video analysis process. In the same region all points collected with GPS technologies were extracted (Fig. 6).

The next step regarded the objects detection activity, which is a sequence of operations aimed at identifying pixels associated with the objects of interest (“foreground” pixels) that are interacting in the scene, separating them from the pixels representing the “background” of the scene. The objects detection algorithm determines the pixels which vary with respect to a frame that represents the background. In the simplest case, the background frame (“reference frame” or “background”) is represented by an image acquired in a previous instant of time in which objects were not present in the scene.

The applicability of this approach to real contexts may be very limited, since it does not take account of changes of the scene in lighting or for a recurrent change due to movement of the vegetation (i.e. trees’ branches moving because of wind). In order to consider all variations that can occur in the background, the most advanced video analysis systems perform sophisticated modeling background techniques able to update the representation of the reference frame automatically to incorporate the changes. However, carrying out this task, the algorithm must be sufficiently “intelligent” to ensure that a vehicle to be tracked, which remains stationary in the scene for a long time (for example a vehicle stopped at a traffic light), is not incorporated into the background.

The presence of shadows and reflections make objects look different in relation to reality and make the object identification procedure even more complex. The camouflage effect (certain vehicles may have similar color characteristics to the road surface) implies that large portions of objects may not be detected.

The video stabilization, the conversion to grayscale space and the Gaussian-blurring filter applied to each frame minimized all errors in the foreground detection, cleaning up each frame of video processing.

This procedure allowed the identification and tracking of all the moving objects within the Region Of Interest; however, it was not able to identify a specific vehicle. In order to identify the target vehicle equipped with GPS tracker and to allow the comparison of measurements provided by different technologies (GPS and UAV aerial photography), a technique based on Haar-like features was applied. The Haar classifier technique allows the addition or subtraction of rectangular regions before processing the result.

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

The peculiarity of the algorithm based on Haar-like technique features consists in being able to adapt the method for the recognition of multiple objects, even in the same sequence, as a result of an appropriate training performed through a large set of samples.

The training of the "cascade" OpenCV library requires indeed a set of positive images, which must contain the object that the procedure has to identify; a set of negative images which do not contain the object is also required, as the negative images represent the background. A large number of positive and negative images are required in order to increase the accuracy of the classifier.

An important parameter that must be taken into account is the number of stages required by the algorithm to generate the classifier. Undertraining the classifier with a small number of stages generates a large number of false positives because the time used in the procedure is not sufficient to determine positive images. If too many stages are set for the training step, it can determine a bad result as positive objects in the picture can be considered as negatives. The algorithm generates a box including the instrumented vehicle. The coordinates of the pixels corresponding to the centroid of the box and coinciding with the target placed on the roof of the vehicle were stored in a dataset.

The dataset includes latitude, longitude and a time instant used for synchronizing UAV and GPS outputs.

The synchronization process begins when the UAV GPS antenna receives from at least four satellites for position and time calculation, then it generates an event that initiates the video capture from UAV camcorder and associates the GPS timestamp with the first frame. UAV dataset and the instrumented vehicle dataset can be directly compared because all GPS devices are synchronized via the satellite signal.

In order to reduce the loss of information caused by the frequency offset of the two different systems, the higher frequency source data was resampled by removing all instances mismatching the lower frequency source data used as a reference.

Case studies and results analysis

Usually, the video image processing technique using UAV applications are employed in a restricted traffic area (parking areas, test circuits, etc.) and does not represent real traffic conditions. In order to test UAV applications for estimating ordinary traffic conditions, the authors chose real urban road sections. The above-mentioned methodology was applied to a set of experiments performed in the city of Milazzo (Messina, Sicily).

Two different sites were analyzed:

- a great urban roundabout at the intersection of the "Asse Viario" with De Gasperi road (Fig. 7a);
- a compact urban roundabout along the road SS 113 (Fig. 7b).

Concerning the video processing stage, a set of 270 images was created manually clipping the marked vehicle in each frame of the acquired video. For each clipped image, using an OpenCV module to generate more positive images, 20 samples were generated by distorting the original image and applying a random background. As the tool does not allow the distortion for multiple images to be applied, a python script was used to generate samples in batch and merge the output files in the binary format required by OpenCV.

In particular, a set of 5400 positive images and 1000 negative images was used, attempting the best configuration of the algorithm parameters in order to calibrate the procedure and reducing unwanted behavior caused by overtraining or under-

Fig. 7. (a) Roundabout at the end of Asse Viario, (b) roundabout along the SS113.

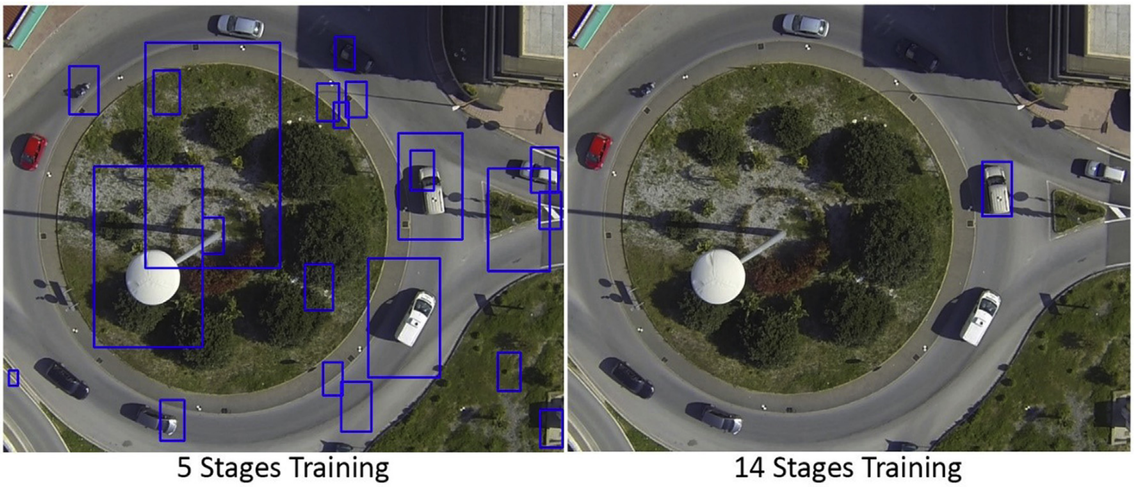

Fig. 8. Stage number parameter tuning.

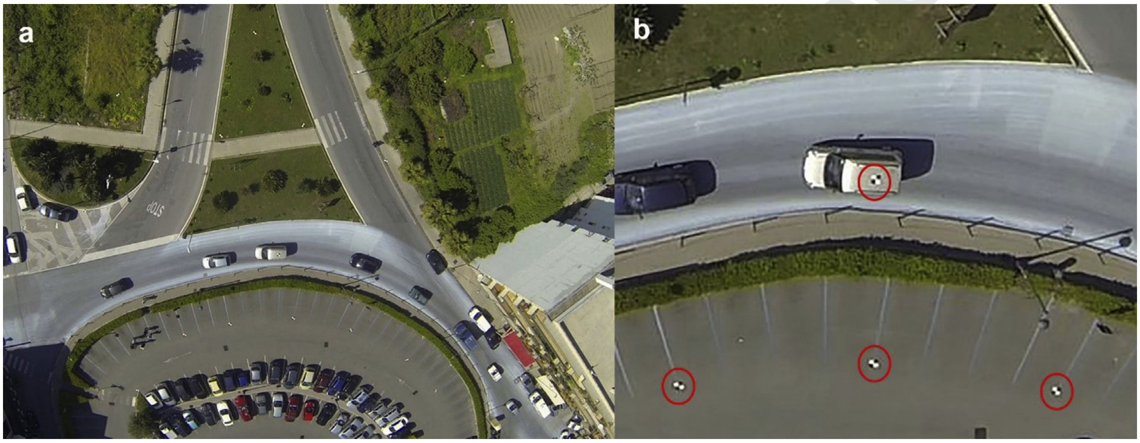

Fig. 9. (a) Sample video frame, (b) position of some control points on the ground and on the vehicle.

training of the classifier. The training procedure in 14 stages produced the best results. A sample of the procedure is illustrated in Fig. 8.

For the first case study, the UAV flew over the roundabout at an altitude of 60 m in hovering (zero speed and constant altitude) from a unique nadir acquisition point. Before starting the combined survey, twenty Ground Control Points (GCP) were positioned and an additional control point was installed on the roof of the probe vehicle. Fig. 9a shows a sample video frame captured during this experiment, while Fig. 9b highlights the position of some control points on the ground and on the car roof.

Five flights were undertaken with the remote-controlled drone for a total of 29′58″ of Full HD Video (1920 $\times$ 1080). At the same time, the V-Box unit recorded speeds and trajectories of the probe vehicle traveling along the road. During the video data acquisition, the instrumented vehicle generated about 15,000 GPS trajectory points falling into the considered Region Of Interest (ROI) of the analyzed frames (Table 3).

Table 3
UAV flights and total number of analyzed frames for case study 1.

ID flight	Total flight time (sec)	Useful flight time (sec)	Probe vehicle transits	Analyzed frames
1	5′ 22″	2′ 53″	4	3979
2	5′ 35″	3′ 37″	5	4991
3	6′ 44″	4′ 26″	7	6118
4	5′ 40″	3′ 39″	6	5037
5	6′ 37″	4′ 38″	8	6394

The spatial accuracy was evaluated in terms of Normalized Root Mean Square Error (NRMSE) between GPS and UAV positions of the instrumented vehicle over all valid tracks.

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{N}[P_{GPSi} - P_{UAVi}]^2}{N}} \tag{1}$$

where

P_{GPSi} = position of the V-Box GPS receiver;

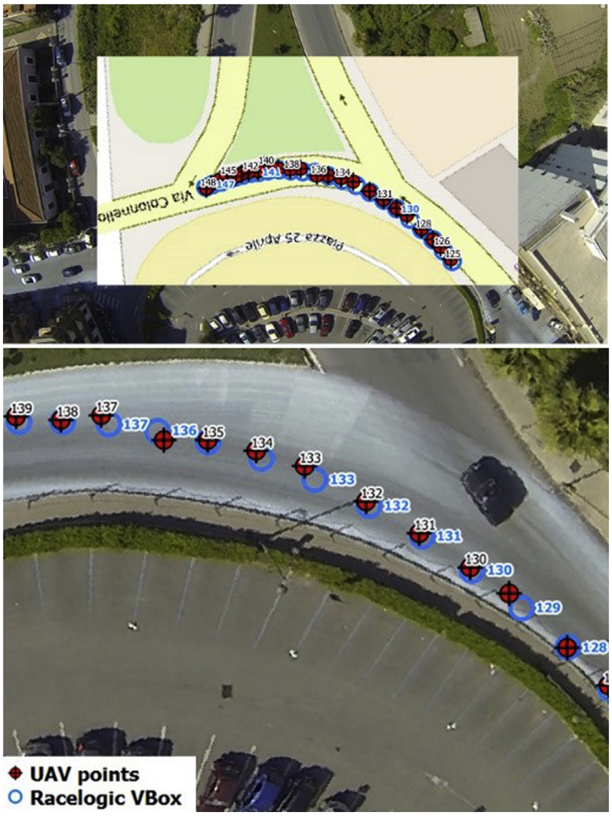

Fig. 10. UAV identified points and V-Box GPS points (case study 1).

Fig. 11. Speed profile comparison for a sample trajectory between UAV and Video V-Box data for case study 1.

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

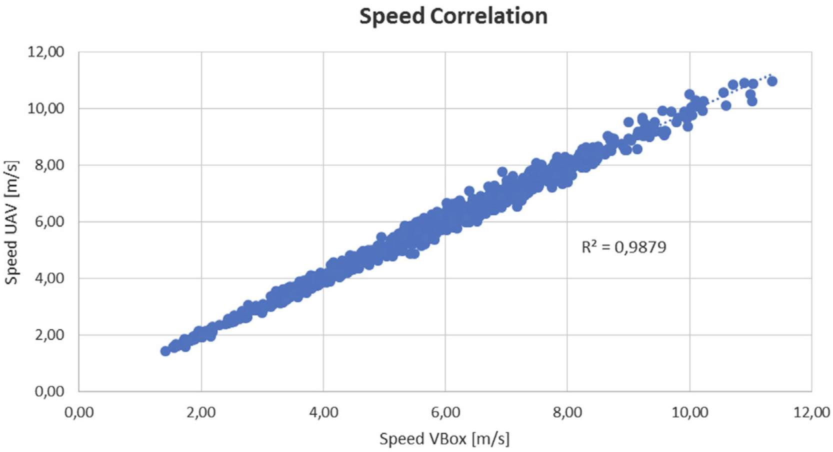

Fig. 12. Speed correlation between UAV and video V-Box data for case study 1.

P_{UAVi} = centroid position of the box from UAV video processing;
N = number of compared position points.

Fig. 10 shows the UAV and V-Box GPS points for a sample trajectory of the instrumented vehicle for this case study.

The Normalized Root Mean Square Error value of 0.55 m obtained for this experiment can be considered a good result for the application purpose. This accuracy in positioning allows using data onto behavioral analysis of single drivers and for determining the level of interaction among vehicles in the traffic stream.

Moreover, speed values from GPS and UAV were compared. GPS speed values are obtained directly from the GPS receivers by the Doppler effect, while speed values coming from the video processing are calculated from the space–time diagrams of the target vehicle. A sample of the speed profile comparison is illustrated in Fig. 11 for a specific trajectory.

The coefficient of correlation for speed reached a high value of 0.98 as shown in Fig. 12. Moreover, the level of accuracy in speed evaluation was expressed also in terms of Root Mean Square Percentage Error reaching the value of 3.57%. For ordinary urban traffic conditions this is an indication of very good accuracy.

Fig. 13. Sample video frame recorded by the drone.

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

Table 4
UAV flights and total number of analyzed frames for case study 2.

ID flight	Total flight time (sec)	Useful flight time (sec)	Probe vehicle transits	Analyzed frames
1	6′ 31″	4′ 03″	7	5589
2	6′ 32″	3′ 44″	6	5152
3	5′ 56″	3′ 43″	6	5129

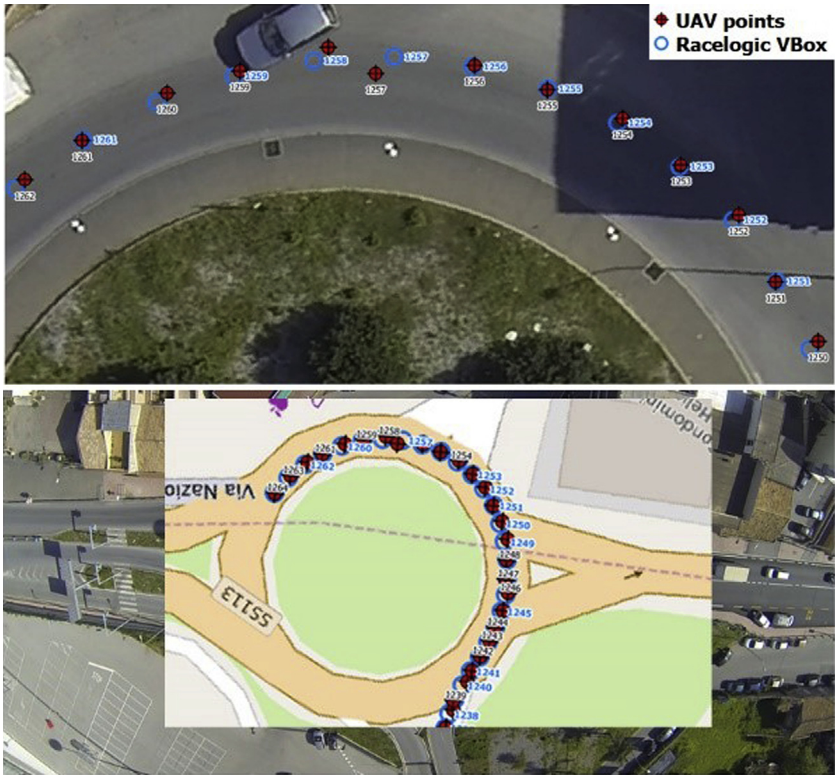

Fig. 14. UAV identified points and V-Box GPS points (case study 2).

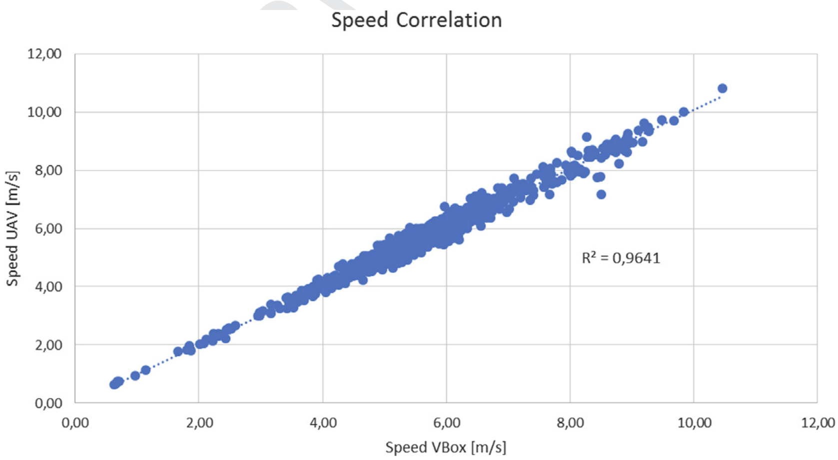

Fig. 15. Speed correlation between UAV and Video V-Box data for case study 2.

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

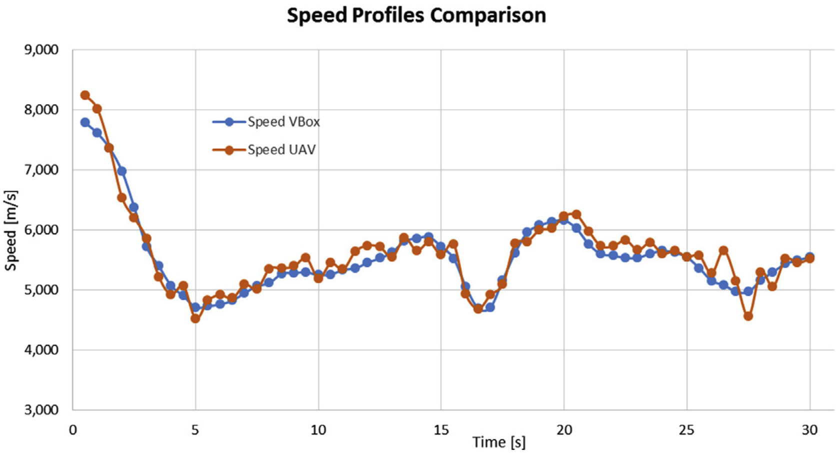

Fig. 16. Speed profile comparison for a sample trajectory between UAV and video V-Box data for case study 2.

Also for the second case study, UAV was used to track the instrumented vehicle trajectories. The compact roundabout is located along the SS 113 road which is a major street with two separated carriageway, each one with two lanes. A preliminary study was done to define the operative parameters of the instrumentations (drone altitude, GPS acquisition time, etc.). Three flights were performed for a total period of 18′59″. The flight plane had a unique nadir acquisition point with the drone hovering at an altitude of 60 m. Fig. 13 shows a sample video frame recorded by the drone (Table 4).

The accuracy measure for position estimates, evaluated in terms of Normalized Root Mean Square Error (NRMSE) between GPS and UAV positions of the instrumented vehicle, was 0.54 m. Comparing these results with those obtained in the first case study (0.55 m) NRMSE indicates that errors could be systematic. An in-depth analysis should be performed on other sites with different road geometric features to confirm this thesis.

Fig. 14 shows the UAV and the V-Box GPS points for a sample trajectory of the instrumented vehicle for this case study.

The assessment of the accuracy in speed evaluation suggests a good level of precision with a coefficient of correlation of 0.96 as shown in Fig. 15. The Root Mean Square Percentage Error is 3.96%. A sample of the speed profile comparison is illustrated in Fig. 16 for a specific trajectory.

Summary and conclusions

In this paper the authors presented a methodology to extract vehicle trajectories and speeds from Unmanned Aerial Vehicles (UAV) video processing. These individual vehicle trajectories were compared with the same ones obtained from a high frequency GPS (Global Positioning System) receiver mounted on board, in order to prove the usefulness and the accuracy of UAV in acquiring reliable traffic data. For this purpose an instrumented vehicle, equipped with a high precision GPS, was driven interacting with the other vehicles of the traffic stream.

In the first part, the paper introduced a video capture experiment performed by UAV in two roundabout case studies. A set of Ground Control Points (GCP) was used to improve the accuracy of the video processing stage. Owing to several disturbances that affected the video (e.g. barrel distortion and noise caused by UAV vibrations), a video stabilization, a conversion to grayscale space and the Gaussian-blurring filters were applied to each frame of the video. Then, in order to separate the target vehicle trajectories from the other ones, the authors applied the Haar classifier approach.

After the video processing stage, a Region Of Interest (ROI) was defined where all points of the target vehicle trajectories collected through the GPS receiver were mapped. Comparing these trajectories monitored by UAV with the GPS traces used as a benchmark, the analysis showed that:

- the Normalized Root Mean Square Error in positioning ranging from 0.54 to 0.55 m;
- the speed profiles present good coefficients of correlation, 0.96 and 0.98, for both case studies;
- the Root Mean Square Percentage Error in speed evaluation is equal to 3.96% and to 3.57%, for case study 1 and 2 respectively.

Therefore, these results demonstrate the importance and the usefulness of UAV in the vehicle trajectories extraction. However, an in-depth analysis is still needed to affirm this adopted procedure can be transferable to all ''noninstrumented'' vehicles transiting on the analyzed sites.

In future studies, attention will be focused on the possibility of calibrating simulation models with a high level of detail by using spatial information acquired from a UAV. At the same time, the observed level of accuracy in speed estimation can be used for safety assessment where the differential speed between a pair of vehicles is the most important factor. Furthermore,

Please cite this article in press as: Guido, G., et al. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology (2017), http://dx.doi.org/10.1016/j.ijtst.2016.12.001

it will be important to consider other additional intrinsic camera parameters (including, for example, lens distortion) to better match the ground coordinate systems. This would help to remove most of the errors due to uncontrolled movements or vibrations of the vehicle arising from wind.

Acknowledgements

The authors would like to thank the staff of E. Lab S.r.l., an Academic Spin-Off of the University of Palermo, for the UAV instrumentation and for its cooperation during the survey stages.

References

Apeltauer, J., Babinec, A., Herman, D., Apeltauer, T., 2015. Automatic vehicle trajectory extraction for traffic analysis from aerial video data, the international archives of the photogrammetry. Remote Sens. Spat. Inf. Sci. 43 (W2), 9–15.
Astarita, V., Bertini, R., d'Elia, S., Guido, G., 2006. Motorway traffic parameter estimation from mobile phone counts. Eur. J. Oper. Res. 175 (3), 1435–1446.
Bai, S., Oh, J., Jung, J., 2013. Context awareness beacon scheduling scheme for congestion control in vehicle to vehicle safety communication. Ad Hoc Netw. 11, 2049–2058.
Baker, K., Sullivan, G., 1992. Performance assessment of model-based tracking. In: Proceedings of the IEEE Workshop on Applications of Computer Vision, Palm Springs CA, pp. 28–35.
Barmpounakis, E.N., Vlahogianni, E.I., Golias, J.C., Extracting kinematic characteristics from unmanned aerial vehicles. In: Proceedings of the Transportation Research Board 95th Annual Meeting, Washington D.C.
Bradski, G., Kaehler, A., 2008. Learning OpenCV: Computer Vision with the OpenCV library, O'Reilly Series. O'Reilly Media, Incorporation, Sebastopol, CA.
Braut, V., Culjak, M., Vukotic, V., Segvic, S., Sevrovic, M., Gold, H., 2012. Estimating OD matrices at intersections in airborne video – a pilot study, MIPRO. In: Proceedings of the 35th International Convention, pp. 977–982.
Chachich, A.C., Pau, A., Barber, A., Kennedy, K., Olejniczak, E., Hackney, J., Sun, Q., Mireles, E., 1996. Traffic sensor using a color vision method. In: Proceedings of SPIE Transportation Sensors and Controls: Collision Avoidance Traffic Management and ITS 2902, pp. 156–164.
Cheng, H.Y., Weng, C.C., Chen, Y.Y., 2012. Vehicle detection in aerial surveillance using dynamic bayesian networks. IEEE Trans. Image Process. 21 (4), 2152–2159.
Choi, K., Chung, Y., 2001. Travel time estimation algorithm using GPS probe and loop detectors data fusion. In: Proceeding of the Transportation Research Board 80th Annual Meeting, Washington D.C.
Coifman, B., Dhoorjaty, S., Lee, Z.H., 2003. Estimating median velocity instead of mean velocity at single loop detectors. Transp. Res. Part C 11C (3–4), 211–222.
Coifman, B., McCord, M., Mishalani, R.G., Redmill, K., 2004. Surface transportation surveillance from unmanned aerial vehicles. In: Proceedings of the Transportation Research Board 83rd Annual Meeting, Washington D.C.
Coifman, B., McCord, M., Mishalani, R.G., 2006. Traffic flow data extracted from imagery collected using a micro unmanned aerial vehicle. In: Applications of Advanced Technology in Transportation the 9th International Conference, pp. 298–303.
Doğan, S., Serhan, T.M., Sitki, K., 2010. Real time speed estimation of moving vehicles from side view images from an uncalibrated video camera. Sensors 10, 4805–4824.
Doherty, P., Granlund, G., Kuchcinski, K., Sandewall, E., Nordberg, K., Skarman, E., Wiklund, J., 2000. The WITAS unmanned aerial vehicle project. In: Proceedings of the 14th European Conference on Artificial Intelligence, pp. 747–755.
Federal Highway Administration, 2007. A new look at sensors. Public Roads 71 (3), 32–39 (US Department of Transportation).
Grabner, H., Nguyen, T.T., Gruber, B., Bischof, H., 2008. On-line boosting-based car detection from aerial images. ISPRS J. Photogramm. Remote Sens. 63, 382–396.
Granlund, G., Nordberg, K., Wiklund, J., Doherty, P., Skarman, E., Sandewal, E., 2000. Witas: an intelligent autonomous aircraft using active vision. In: Proceedings of the UAV 2000 International Technical Conference and Exhibition.
Guido, G., Vitale, A., Astarita, V., Saccomanno, F.F., Giofrè, V.P., Gallelli, V., 2012. Estimation of safety performance measures from smartphone sensors. Procedia Social Behav. Sci. 54, 1095–1103.
Guido, G., Vitale, A., Saccomanno, F.F., Festa, D.C., Rogano, D., Gallelli, V., 2014. Treating uncertainty in the estimation of speed from smartphone traffic probes. Transp. Res. Part C 47 (1), 100–112.
Haar, A., 1910. Zur Theorie der orthogonalen Funktionensysteme. Math. Ann. 69 (3), 331–371.
Handel, P., Ohlsson, J., Ohlsson, M., Skog, I., Nygren, E., 2014. Smartphone based measurement systems for road vehicle traffic monitoring and usage based insurance. IEEE Syst. J. 8 (4), 1238–1248.
Haoui, A., Kavaler, R., Varaiya, P., 2008. Wireless magnetic sensors for traffic surveillance. Transp. Res. Part C 16 (3), 294–306.
Herrera, J.C., Work, D.B., Herring, R., Banc, X., Jacobson, Q., Bayen, A., 2010. Infrastructure-based technique for traffic data acquirement. Transp. Res. Part C 18 (4), 568–583.
Hinza, S., Stilla, U., 2006. Car detection in aerial thermal images by local and global evidence accumulation. Pattern Recognit. Lett. 27 (4), 308–315.
Ho, T.J., Chung, M.J., 2016. Information-aided smart schemes for vehicle flow detection enhancements of traffic microwave radar detectors. Appl. Sci. 6 (7), 196.
Huang, M.C., Yen, S.H., 2004. A real-time and color-based computer vision for traffic monitoring system. In: IEEE International Conference on Multimedia and Expo, vol. 3, Taipei Taiwan, pp. 2119–2122.
Kamijo, S., Ikeuchi, K., Sakauchi, M., 2001. Vehicle tracking in low-angle and front view images based on spatio-temporal Markov random fields. In: Proceedings of the 8th World Congress on Intelligent Transportation Systems.
Kanhere, N.K., Birchfield, S., 2008. Real-time incremental segmentation and tracking of vehicles at low camera angles using stable features. IEEE Trans. Intell. Transp. Syst. 9 (1), 148–160.
Kim, S.W.E., 1998. Performance comparison of loop/piezo and ultrasonic sensor-based traffic detection systems for collecting individual vehicle information. In: Proceedings of the 5th World Congress on Intelligent Transport Systems, 4083.
Koller, D., Daniilidis, K., Nagel, H.H., 1993. Model-based object tracking in monocular image sequences of road traffic scenes. Int. J. Comput. Vis. 10 (3), 257–281.
Koller, D., Weber, J., Malik, J., 1994. Robust Multiple Car Tracking With Occlusion Reasoning. ECCV, Stockholm Sweden, pp. 189–196.
Kwong, K., Kavaler, R., Rajagopal, R., Varaiya, P., 2009. Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors. Transp. Res. Part C 17 (6), 586–606.
Lai, A.H.S., Fung, G.S.K., Yung, N.H.C., 2001. Vehicle type classification from visual-based dimension estimation. In: Proceedings of the IEEE Intelligent Transportation Systems Conference, pp. 201–206.
Lee, J., Zhong, Z., Kim, K., Dimitrijevic, B., Du, B., Gutesa, S., 2015. Examining the applicability of small quadcopter drone for traffic surveillance and roadway incident monitoring. In: Proceedings of the Transportation Research Board 94th Annual Meeting, Washington D.C.
Li, Z., Yang, X., 2006. Application of cement-based piezoelectric sensors for monitoring traffic flows. J. Transp. Eng. 132 (7), 565–573.
Lienhart, R., Maydt, J., 2002. An extended set of Haar-like features for rapid object detection. Proc. Int. Conf. Image Process., 900–903

McGowen, P., Sanderson, M., 2011. Accuracy of pneumatic road tube counters. In: 2011 Western District Annual Meeting. Institute of Transportation Engineers.

Nooralahiyan, A.Y., Kirby, H.R., McKeown, D., 1998. Vehicle classification by acoustic signature. Math. Comput. Modell. 27 (9–11), 205–214.

Oh, J., Min, J., Kim, M., Cho, H., 2009. Development of an automatic traffic conflict detection system based on image tracking technology. In: Proceedings of the Transportation Research Board 88th Annual Meeting, Washington D.C.

OpenCV Library 2015. Open Source Computer Vision Library. In: <http://docs.opencv.org/3.1.0/>.

Papageorgiou, C., Oren, M., Poggio, T., 1998. A general framework for object detection. In: Proceedings of the Sixth International Conference on Computer Vision, pp. 555–562.

Polk, A.E., Pietrzyk, M.C., 1995. The Miami method: using automatic vehicle location (AVL) for measurement of roadway level-of-service. In: Proceedings of the Annual Meeting of ITS America, vol. 1. Intelligent Transportation Society of America, Washington D.C..

Puri, A., 2005. A Survey of Unmanned Aerial Vehicles (UAV) for Traffic Surveillance. Department of Computer Science and Engineering, University of South Florida, pp. 1–29.

Puri, A., Valavanis, K., Kontitsis, M., 2007. Statistical profile generation for traffic monitoring using real-time UAV based video data. In: Mediterranean Conference on Control & Automation, pp. 1–6.

Rad, R., Jamzad, M., 2005. Real time classification and tracking of multiple vehicles in highways. Pattern Recognit. Lett. 26 (10), 1597–1607.

Rango, A., Laliberte, A., Steele, C., Herrick, J.E., Bestelmeyer, B., Schmugge, T., Roanhorse, A., Jenkins, V., 2006. Using unmanned aerial vehicles for rangelands: current applications and potential. Environ. Pract. 8, 159–168.

Salvo, G., Caruso, L., Scordo, A., 2014. Urban traffic analysis through an UAV. Procedia Social Behav. Sci. 111, 1083–1091.

Sanguesa, J.A., Fogue, M., Garrido, P., Martinez, F.J., Cano, J.C., Calafate, C.T., Manzoni, P., 2013. An infrastructureless approach to estimate vehicular density in urban environments. Sensory 13 (2), 2399–2418.

Schlosser, C., Reitberger, J., Hinz, S., 2003. Automatic car detection in high resolution urban scenes based on an adaptive 3D-model. In: EEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Berlin, pp. 98–107.

Smalley, D.O., Hickman, D.R., McCasland, W.R., 1996. Design and Implementation of Automatic Vehicle Identification Technologies for Traffic Monitoring in Houston, Texas, Draft Report TX-97-1958-2F. Texas Transportation Institute, College Station, Texas.

Song, K.T., Chen, C.H., Huang, C.H.C., 2004. Design and experimental study of an ultrasonic sensor system for lateral collision avoidance at low speeds. In: IEEE Intelligent Vehicles Symposium, pp. 647–652.

Sumner, R., Smith, R., Kennedy, J., Robinson, J., 1994. Cellular based traffic surveillance – the Washington, D.C. Area operational test. In: Proceeding of the IVHS America Annual Meeting, Vol. 2, Washington D.C.

Sun, C., Ritchie, S., 1999. Individual vehicle speed estimation using single loop inductive waveforms. J. Transp. Eng. 125 (6), 531–538.

Tyagi, V., Kalyanaraman, S., Krishnapuram, R., 2012. Vehicular traffic density state estimation based on cumulative road acoustics. IEEE Trans. Intell. Transp. Syst. 13 (3), 1156–1166.

Vaidya, N., Higgins, L.L., Turnbull, K.F., 1996. An evaluation of the accuracy of a radio trilateration automatic vehicle location system. In: Proceedings of the Annual Meeting of ITS America. Intelligent Transportation Society of America, Washington D.C..

Viola, P., Jones, M., Rapid object detection using a boosted cascade of simple features. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 511–518.

Yim, Y.B., Cayford, R., 2001. Investigation of Vehicles as Probes Using Global Positioning System and Cellular Phone Tracking: Field Operational Test. University of Berkeley, Berkley California (California PATH Working Paper UCB-ITS-PWP-2001-9).

Zehang, S., Bebis, G., Miller, R., 2002. Improving the performance of on-road vehicle detection by combining gabor and wavelet features. In: Proceedings of the IEEE International Conference on Intelligent Transportation Systems.

Zhang, G., Avery, R.P., Wang, Y., 1993. Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras. Transp. Res. Rec. 2007, 138–147.

Zwahlen, H.T., Russ, A., Oner, E., Parthasarathy, M., 2005. Evaluation of microwave radar trailers for nonintrusive traffic measurements. Transp. Res. Rec. 1917 (1), 127–140.