Research Article

Traffic Data Collection for Floating Car Data Enhancement in V2I Networks

D. F. Llorca, M. A. Sotelo, S. Sánchez, M. Ocaña, J. M. Rodríguez-Ascariz, and M. A. García-Garrido

Department of Electronics, University of Alcalá, Ctra. N-II Km. 33, Alcalá de Henares, C.P. 28871 Madrid, Spain

Correspondence should be addressed to D. F. Llorca, llorca@depeca.uah.es

This paper presents a complete vision-based vehicle detection system for floating car data (FCD) enhancement in the context of vehicular ad hoc networks (VANETs). Three cameras (*side-, forward-* and *rear*-looking cameras) are installed onboard a vehicle in a fleet of public buses. Thus, a more representative local description of the traffic conditions (extended FCD) can be obtained. Specifically, the vision modules detect the number of vehicles contained in the local area of the host vehicle (*traffic load*) and their relative velocities. Absolute velocities (*average road speed*) and global positioning are obtained after combining the outputs provided by the vision modules with the data supplied by the CAN Bus and the GPS sensor. This information is transmitted by means of a GPRS/UMTS data connection to a central unit which merges the extended FCD in order to maintain an updated map of the traffic conditions (*traffic load* and *average road speed*). The presented experiments are promising in terms of detection performance and computational costs. However, significant effort is further necessary before deploying a system for large-scale real applications.

1. Introduction

Floating car data (FCD) refers to technology that collects traffic state information from a set of individual vehicles which *float* in the current traffic. Each vehicle, which can be seen as a moving sensor that operates in a distributed network, is equipped with positioning (GPS) and communication (GSM, GPRS, UMTS, etc.) systems, transmitting its location, speed, and direction to a central control unit that integrates the information provided by each one of the vehicles.

FCD systems are being increasingly used in a variety of important applications since they overcome the limitations of fixed traffic monitoring technologies (installation and maintenance costs, lack of flexibility, static nature of the information, etc.). We refer to [1] for general background concerning the most representative FCD activities in Japan, Europe, and the United States.

FCD can be used by the public sector to collect road traffic statistics and to carry out real-time road traffic control. The information provided by FCD systems can be supplied to individual drivers via dynamic message signs, PDA devices, satellite navigation systems, or mobile phones, including dynamic rerouting information. Thus, drivers would be able to make more informed choices, spending less time in congested traffic. In addition, the knowledge of the current traffic situation can be also used to estimate time of arrival of a fleet of public transport vehicles and, furthermore, to plan and coordinate the movements of the fleet (fleet management) so that driving assignments can be carried out more efficiently. Besides previous applications, the use of FCD entails environmental benefits since it can be used to reduce fuel consumption and emissions.

The basic data provided by FCD systems (vehicle location, speed, and direction) can be enriched using new onboard sensors (ambient temperature, humidity and light, windshield wiper status, fog light status, fuel consumption, emissions, tire pressure, suspension, emergency brake, etc.) which are centralized by means of the controller-area-network (CAN) bus. Such data can be exploited to extend the information horizon including traffic, weather, road management, and safety applications [1]. In addition,

computer vision systems can be included in order to improve the automatic detection of potentially interesting events and to document them by sending extended data [2].

In order to provide ubiquitous coverage of the entire road network, a minimum representation of the total passenger car fleet has to be used, since each moving sensor (each vehicle) only supplies information about its status. The fact that everyday road users have to be asked to share information regarding their movements and speeds arises privacy issues that have to be addressed. Many potential road travellers may be reluctant to join FCD projects because of violations of their privacy due to permanent traceability or possible liability in case of speed limit violations. Thus, the fundamental concept for FCD systems calls for no identification information to be sent with the basic data, which can be easily implemented from a technical perspective. For example, in [3] a general method for anonymization of FCD by deriving pseudonyms for trips is presented.

Another approach consists of using the information supplied by a specific fleet of vehicles, rather than information coming from individual road users. Taxis or public transport buses can be used due to the extended periods of time they spend on the urban road network. Although taxis and buses provide a major source of innercity traffic information because of the time they spend mobile, they have limitations. Problems arise if the taxi drivers, through detailed knowledge of the local road network, take steps to avoid congested areas which will not be reported [4]. Traffic load perception may be lower than the actual one if reserved taxis or buses lanes are used. On the contrary, privacy issues are not as critical as before, especially when using a fleet of public transport buses.

This paper presents a complete vision-based vehicle detection system onboard a fleet of public transport buses with the aim of improving the data collected in FCD applications. The proposed system has been developed in the framework of the GUIADE project. Three cameras covering the local environment of the vehicle are used: *forward-rear-* and *side*-looking cameras. The system obtains under certain constraints, such as good weather and daytime conditions, the number of vehicles in the local range of the bus as well as their relative position and velocity. This information is combined with the data provided by regular FCD systems (global location, speed, and direction), obtaining a more detailed description of the local traffic load and the average speed. The communication system between the vehicles and the central control unit is based on wireless technology via GPRS/UMTS cellular protocols. Finally, the central unit integrates the data collected by the fleet in order to generate updated traffic status maps.

The remainder of this paper is organized as follows: the description of the system including the wireless communication scheme is summarized in Section 2. Section 3 describes the vision-based vehicle detection system as well as the spatial and temporal integration of the collected data. Experimental results that validate the proposed approach are presented in Section 4. Finally, conclusions and future works are discussed in Section 5.

2. System Descripction

The proposed FCD architecture can be seen in Figure 1. Floating car data is supplied by a fleet of public transport buses which corresponds to an inner-city bus line. Each vehicle is equipped with a global positioning system (GPS), wireless communication interfaces (GPRS/UMTS and WLAN IEEE 802.11) and a complete vision-based vehicle detection system.

The vehicle-to-infrastructure (V2I) communication system is based on the geographic coverage provided by cellular networks. General packet radio service (GPRS) and universal mobile telecommunications system (UMTS) are used to connect each vehicle with the central control unit. Each vehicle provides information that can be divided in three main groups.

(1) *Standard FCD information*: vehicle identifier (2 bytes), timestamp (11 bytes), GPS position (8 bytes), speed (2 bytes), and direction (2 bytes).

(2) *Vehicle status information*: ambient temperature (2 bytes), humidity (2 bytes), light (2 bytes), windshield wiper status (1 byte), fog light status (1 byte), fuel consumption (4 bytes), and emissions (4 bytes).

(3) *Extended FCD information*: globally referenced average traffic load (2 bytes) and average road speed for a measured segment travel time (2 bytes).

As can be observed, the total message size per vehicle is 45 bytes. The extended FCD information is supplied to the central unit at a frequency of 1 Hz. Accordingly, the bandwidth currently demanded by vehicular communication in the communication channel, that is, the vehicle throughput, is 360 bps without overheads. This value can be considered negligible taking into account the available bandwidth and the proposed FCD architecture.

The central control unit integrates the information provided by each one of the vehicles in order to compute updated traffic and weather maps which will be used for fleet management tasks as well as to estimate the time of arrival.

The vehicle-to-vehicle (V2V) communication system is defined as a *backup* communication system based on a wireless-fidelity (WiFi) IEEE 802.11a/b/g interface. In situations where the cellular network is not working, in-range vehicles will exchange the most updated information available.

One of the main advantages of the proposed approach is that it does not need to deal with privacy issues since the floating vehicles correspond to a fleet of public transport buses.

3. Vision-Based Traffic Detection System

In this section, we present the main contribution of this work: a complete vision-based traffic detection system which enhances the data supplied by standard FCD systems. The benefits of using computer vision instead of other technologies such as radar-based systems can be summarized as follows. Computer vision systems can compensate for
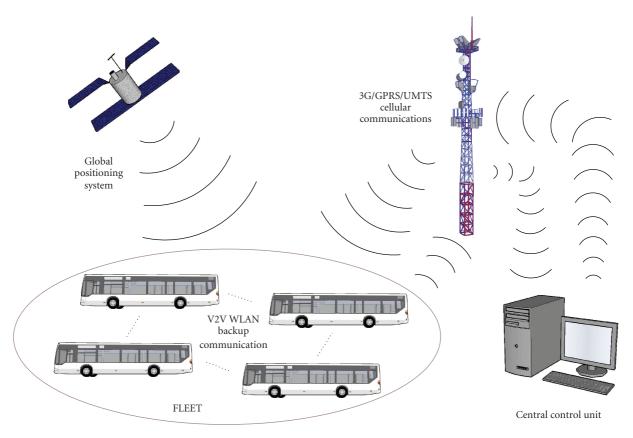
FIGURE 1: Overview of the proposed FCD architecture.

the lower angular resolution of the low-cost radar and the increased appearance of ghost radar targets (guard-rails, railings, lamp posts, reflections, etc.). These false positives are relevant and they cannot simply be ignored. The camera has very good angular resolution and can be used to determine height, width, and lateral speed of the target. Pattern recognition can be used to classify the object and even weakly reflective targets such as pedestrians can be detected. Moreover, the cost of a vision system is significantly lower than the cost saved by using the simpler radar. A vision system, in addition to overcoming cost reduction problems, can contribute to the system features such as road analysis and scene understanding.

Each individual vehicle is equipped with three FireWire cameras (forward-, rear- and side-looking cameras) that cover the local environment of the bus (see Figure 2). A common hardware trigger synchronizes the image acquisition of the three cameras and an onboard PC houses the computer vision software.

Each individual vehicle detection system provides information about the number of detected vehicles and both their relative position and speed. These results are combined with the GPS measurements and the data provided by the CAN bus in order to provide globally referenced traffic information. This scheme is described in Figure 3.

The layers of the proposed architecture of the three vision modules are conceptually the same: *lane detection*, *vehicle candidates selection*, *vehicle recognition,* and *tracking*. The first step of each one of the vision systems consists of reducing the searching space in the image plane in an intelligent manner in order to increase the performance of the vehicle detection module. Accordingly, road lane markings are detected and used as the guidelines that drive the vehicle searching process (see Figure 4). The area contained by the limits of the lanes is scanned in order to find vehicle candidates that are passed on to the vehicle recognition modules. Thus, the rate of false positives is reduced. In case that no lane markings are detected, a basic *region of interest* is used instead covering the front, rear, and side parts of the vehicle. Finally, a tracking stage is implemented using Kalman filtering techniques.

3.1. Lane Detection. An attention mechanism is necessary in order to filter out inappropriate candidate windows based on the lack of distinctive features, such as horizontal edges and vertical symmetrical structures, which are essential characteristics of road vehicles. This has the positive effect of decreasing both the total computation time and the rate of false positive detections. Lane markings are detected using gradient information in combination with a local thresholding method which is adapted to the width of the projected lane markings. Then, clothoid curves are fitted to the detected markings. The algorithm scans up to 25 lines in the candidates searching area, from 2 meters in front of the camera position to the maximum range in order to find the lane marking measurements. The proposed method implements a nonuniform spacing search that
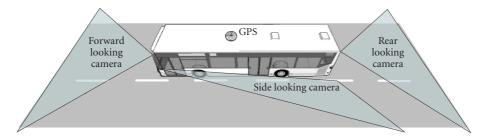
FIGURE 2: Main vehicle sensors: three cameras (forward, rear, and side looking cameras) and a global positioning system.
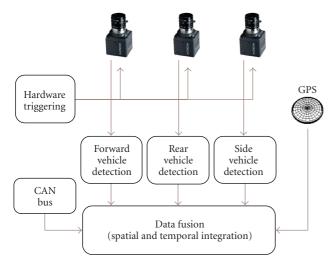

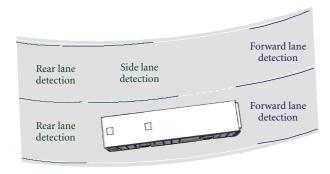FIGURE 3: Stages of the vision-based traffic detection system.

FIGURE 4: Rear, side, and forward lane detection.

reduces certain instabilities in the fitted curve. The final state vector is composed of 6 variables [5] for each lane on the road
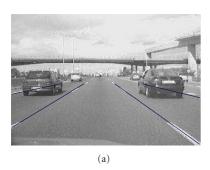
$$x = [c_{oh}, c_{1h}, c_{ov}, c_{1v}, x_o, \theta_o, w_o]^T, \tag{1}$$

where c_{oh} and c_{1h} represent the clothoid horizontal curvature parameters, c_{ov} and c_{1v} stand for the clothoid vertical curvature parameters, while x_o, θ_o, and w_o are the lateral error and orientation error with regard to the centre of the lane and the width of the lane, respectively. The clothoid curves are then estimated based on lane marking measurements using a Kalman filter for each lane.

Apart from the detected road lanes additional virtual lanes have been considered so as to cope with situations in which a vehicle is located between two lanes (e.g., if it is performing a change lane manoeuvre). Virtual lanes provide the necessary overlap between lanes, avoiding both misdetections and double detections caused by the two halves of a vehicle being separately detected as two potential vehicles. A virtual lane is located to provide overlap between two adjoining lanes. Figure 5 provides some examples of lane markings detection in real outdoor scenarios. Detected lanes determine the vehicle searching area and help reduce false positive detections. In case no lane markings are detected by the system, fixed lanes corresponding to a straight road model are assumed instead.

3.2. Side Vehicle Detection. Side vehicle detection module [6] relies on the computation of optical flow. In order to reduce computational time, optical flow is computed only on Canny points in the image. Canny edge pixels are consequently matched and grouped together in order to detect clusters of pixels that can be considered as candidate vehicles in the image. Classical clustering techniques are used to determine groups of pixels, as well as their likelihood to form a single object. Even after pixels clustering, some clusters can still be clearly regarded as belonging to the same real object. A second grouping stage (double-stage) is then carried out among different clusters in order to determine which of them can be further merged into a single blob. For this purpose, simple distance criteria are considered. Two objects that are very close to each other are finally grouped together in the same cluster. The reason for computing a two-stage clustering process relies on the fact that by selecting a small distance parameter in the first stage, interesting information about clusters in the scene can be obtained. Otherwise, using a large distance parameter in the single clustering process, highly gross clusters would have been achieved, losing all information about the granular content of the points that provide optical flow in the image.

The selected clusters constitute the starting point for locating candidate vehicles in the image. For that purpose, the detected positions of clusters are used as a seed point to search for a collection of horizontal edges that could potentially represent the lower part of a car. The candidate is located on the detected horizontal edges that meet certain conditions of entropy and vertical symmetry. Some of the most critical aspects in side vehicle detection are
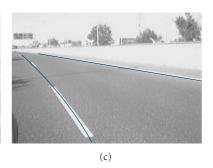
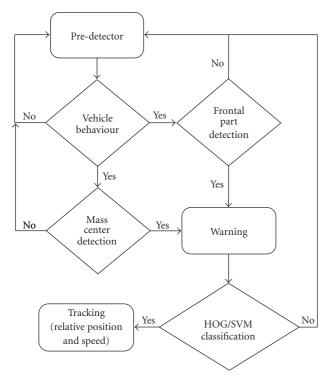FIGURE 5: Vehicle searching area as a result of the lane markings analysis for forward, rear and side modules.

FIGURE 6: Side vehicle detection flow diagram.

with the samples obtained from real road images, and at that point vehicle tracking starts. Tracking is stopped when the vehicle gets out of the image. Sometimes, the shadow of the vehicle remains in the image for a while after the vehicle disappears from the scene, provoking the warning alarm to hold on for 1 or 2 seconds. This is not a problem, however, since the overtaking car is running in parallel with the ego-vehicle during that time although it is out of the image scene. Thus, maintaining the alarm in such cases turns out to be a desirable side effect.

Figure 7 shows an example of blind spot detection in a sequence of images. The indicator depicted in the upper-right part of the figure toggles from green to blue when a vehicle enters the blind spot area (indicated by a green polygon). A blue bounding box depicts the position of the detected vehicle.

3.3. Forward and Rear Vehicle Detection. Forward- and rear-looking vehicle detection systems share the same algorithmic core. The attention mechanism sequentially scans each road lane from the bottom to the maximum range looking for a set of features that might represent a potential vehicle. Firstly, the vehicle contact point is searched by means of the top-hat transformation. This operator allows the detection of contrasted objects on nonuniform backgrounds [9]. There are two different types of top-hat transformations: white hat and black hat. The white hat transformation is defined as the residue between the original image and its opening (∘ operator). The black hat transformation is defined as the residue between the closing (• operator) and the original image. The white and black hat transformations are analytically defined as follows:

$$\text{WH}_T(x, y) = (f - f \circ b)(x, y) \quad \text{White Hat,} \quad (2)$$

$$\text{BH}_T(x, y) = (f \bullet b - f)(x, y) \quad \text{Black Hat.} \quad (3)$$

The opening operator (∘) is defined as the dilation of the erosion and the closing operator (•) is defined as the erosion of the dilation (for more details see [10]). In our case we use the white hat operator (2) since it enhances the boundary between the vehicles and the road [11]. Horizontal contact points are preselected if the number of white top-hat features is greater than a configurable threshold. Then, candidates are preselected if the entropy of Canny points is high enough

subsequently listed: (1) shadows on the asphalt due to lampposts, other artefacts or a large vehicle overtaking the ego-vehicle on the right lane; (2) self-shadow reflected on the asphalt (especially problematic in sharp turns like in round-about points), or self-shadow reflected on road protection fences; (3) robust performance in tunnels; and (4) avoiding false alarms due to vehicles on the third lane.

The flow diagram of the two-stage detection algorithm is depicted in Figure 6. As can be observed, there is a pre-detector that discriminates whether the detected object is behaving like a vehicle or not. If so, the frontal part of the vehicle is located in the region of interest. In addition, the vehicle mass centre is computed. In case the frontal part of the vehicle is properly detected and its mass centre can also be computed, a final warning message is issued. After being located, vehicle candidates are classified by using a linear SVM classifier [7] with HOG features [8] previously trained

(a)   (b)   (c)   (d)
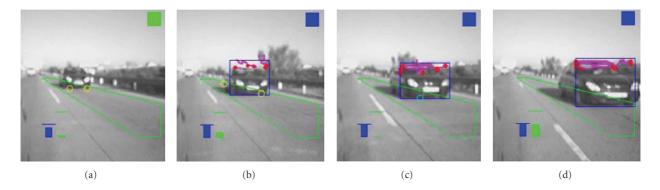
FIGURE 7: Example of the side-vehicle detection module (also called *blind spot* detection) in a sequence of images. The indicator in the upper-right part of the figure toggles from green to blue when a car is detected in the blind spot.

(a)   (b)   (c)
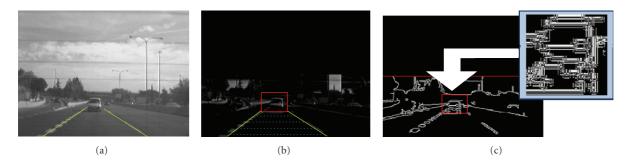
FIGURE 8: From left to right: original image; contact point detection on white top-hat image; candidate preselected with high entropy of Canny points.
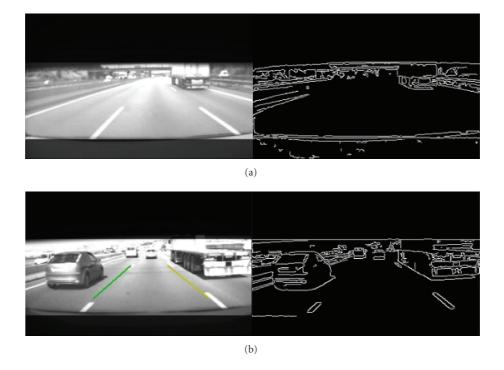

(a)

(b)

FIGURE 9: Canny images after adaptive thresholding.
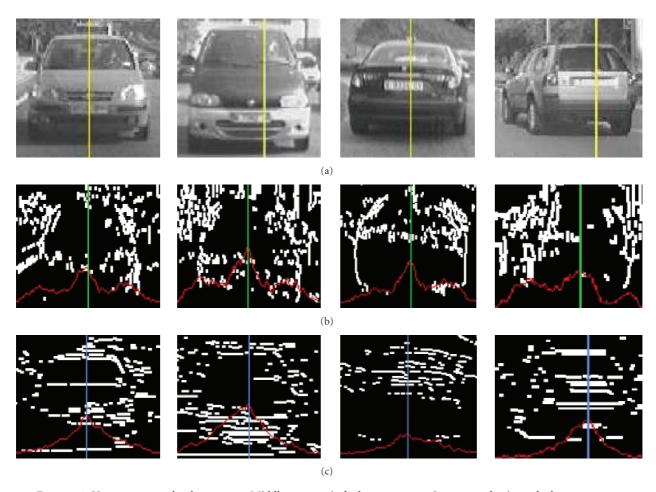
(a)

(b)

(c)

FIGURE 10: Upper row: gray level symmetry; Middle row: vertical edges symmetry; Lower row: horizontal edges symmetry.

(a)																(b)

FIGURE 11: (a) Overlapped candidates. (b) Nonmaximum suppression results.

for a region defined by means of perspective constraints and prior knowledge of target objects (see Figure 8).

Before computing the Canny features, an adaptive thresholding method is applied. This process is based on an iterative algorithm that gradually increases the contrast of the image, and compares the number of Canny points obtained in the contrast increased image with the number of edges obtained in the current image. If the number of Canny features in the actual image is higher than in the contrast increased image the algorithm stops. Otherwise, the contrast is gradually increased and the process resumed. This adaptive thresholding method permits to obtain robust image edges, as depicted in the examples provided in Figure 9.
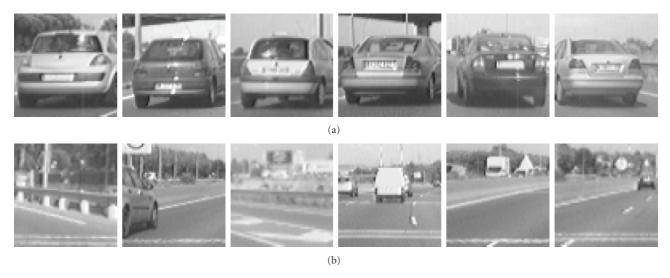
(a)

(b)

Figure 12: Forward data set. (a) positive samples (vehicles). (b) negative samples.

(a)

(b)

Figure 13: Rear data set. (a) positive samples (vehicles). (b) negative samples.

(a)   (b)

Figure 14: Linear SVM with HOG features classification examples: nonvehicle (red) and vehicle (green).

(a)

(b)

(c)

(d)

FIGURE 15: Data association by features matching. (a, b) Harris features on image t. (c, d) matched Harris features on image $t + 1$.

—— Ground truth
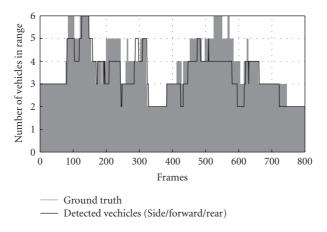—— Detected vechicles (Side/forward/rear)

FIGURE 16: Number of vehicles detected by the three vision modules compared with the manually labeled ground truth in a real sequence.

In a second step, vertical edges (S_v), horizontal edges (S_h), and grey level (S_g) symmetries are obtained, so that, candidates will only pass to the next stage if their symmetries values are greater than a threshold. The vertical and horizontal edges symmetries are computed as listed inAlgorithm 1. The grey level symmetry computation procedure is shown inAlgorithm 2. Some examples of the three types of symmetries are depicted in Figure 10.

Symmetry axes are linearly combined to obtain the final position of the candidate. Finally, a weighted variable is defined as a function of the entropy of Canny points, the three symmetry values and the distance to the host vehicle. We use this variable to apply a nonmaximum suppression process per lane which removes overlapped candidates. An example of this process is shown in Figure 11.
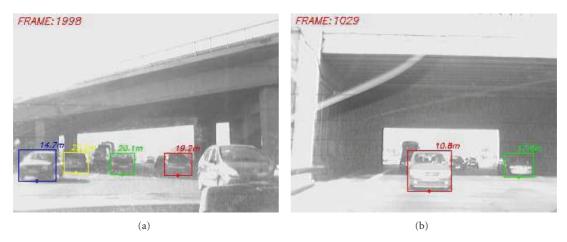
(a)

(b)

FIGURE 17: Examples with strong illumination changes after passing beneath a bridge.

(1) Initialize $\text{Acc}_{0,\ldots,ROI_{\text{WIDTH}}} = 0$
(2) For $i = 0,\ldots,ROI_{\text{HEIGHT}}$
(3)   For each pair of vertical/horizontal edge pixels (x_1, i) and (x_2, i)
(4) $\text{Acc}_{(x_1+x_2)/2}$ + +
(5) $S_{v,h} = \arg\max_i (\text{Acc}_i / S_{v,h,\text{MAX}})$

ALGORITHM 1: Vertical and horizontal edges symmetries computation procedure.

The selected candidates are classified by means of a linear SVM classifier [7], in combination with histograms of oriented gradients features [8]. We have developed and tested two different classifiers depending on the module (forward and rear classifiers). All candidates are resized to a fixed size of 64×64 pixels to facilitate the features extraction process. The rear-SVM classifier is trained with 2000 samples and tested with 1000 samples (1/1 positive/negative ratio) whereas the forward-SVM classifier is trained with 3000 samples and tested with 2000 samples (1/1 positive/negative ratio). Figures 12 and 13 depict some positive and negative samples of the forward and rear training and test data sets, respectively. Figure 14 shows a couple of examples of vehicle detection after linear SVM classification with HOG features.

After detecting consecutively an object classified as vehicle a predefined number of times (empirically set to 3 in this work), data association and tracking stages are triggered. The data association problem is addressed by using feature matching techniques. Harris features are detected and matched between two consecutive frames as depicted in Figure 15.

Tracking is implemented using Kalman filtering techniques [12]. For this purpose, a dynamic state model and a measurement model must be defined. The proposed dynamic state model is simple. Let us consider the state vector x_n, defined as follows:

$$x_n = \left[u, v, w, h, \dot{u}, \dot{v}, \dot{w}, \dot{h} \right]^T. \tag{4}$$

In the state vector x and y are the respective horizontal and vertical image coordinates for the top left corner of every object, and w and h are the respective width and height in the image plane, a dynamical model equation can be written like this

$$x_{n+1} = A \cdot x_n + \omega_n$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ h \\ \dot{u} \\ \dot{v} \\ \dot{w} \\ \dot{h} \end{pmatrix}_n + \omega_n. \tag{5}$$

In the model, Δt is the simple time, A represents the system dynamics matrix and ω_n is the noise associated to the model. Although the definition of A is simple, it proves to be highly effective in practice since the real time operation of the system permits to assure that there will not be great differences in distance for the same vehicle between consecutive frames. The model noise has been modelled as a function of distance and camera resolution. The state model equation is used for prediction in the first step of the Kalman filter. The next step is to define the measurement model. The

(1) For each possible symmetry axis x_i initializes $S_i = 0$
(2)   For $j = 0, ..., ROI_{\text{HEIGHT}}$
(3)   For $k = 0, ..., ROI_{\text{WIDTH}}/2$
(4)   If $\text{abs}(\text{image}_{[j][x_i+k]} - \text{image}_{[j][x_i-k]}) < \Delta$
(5) $S_i {+}{+}$
(6) $S_g = \underset{i}{\arg\max}(S_i/(\text{area}_{ROI}/2))$

ALGORITHM 2: Gray level symmetry computation procedure.

FIGURE 18: Traffic load L_i at every frame in a real sequence.

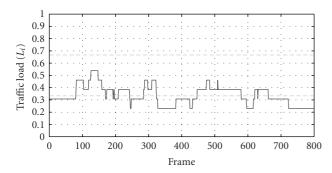FIGURE 20: Average traffic load at every second in a real sequence.

FIGURE 19: Average road speed v_i at every frame in a real sequence.
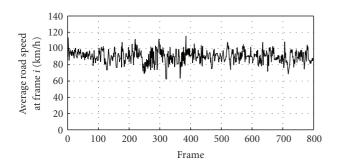

FIGURE 21: Average road speed at every second in a real sequence.

measurement vector is defined as $z_n = [u, v, w, h]^T$. Then, the measurement model equation is established as follows:

$$z_{n+1} = H \cdot x_n + v_n$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ h \\ \dot{u} \\ \dot{v} \\ \dot{w} \\ \dot{h} \end{pmatrix}_n + v_n. \quad (6)$$

In last equation, H represents the measurement matrix and v_n is the noise associated to the measurement process. The purpose of the Kalman filtering is to obtain a more stable position of the detected vehicles. Besides, oscillations in vehicles position due to the unevenness of the road makes v coordinate of the detected vehicles change several pixels up
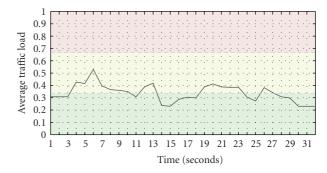
or down. This effect makes the distance detection unstable, so a Kalman filter is necessary for minimizing these kinds of oscillations.

3.4. FCD Integration. As depicted in Figure 3, the FCD integration or Data Fusion module uses three sources of data: the measurements provided by the GPS, the data supplied by the CAN bus, and the output obtained from the three vision-based vehicle detection modules. Whereas the GPS and the CAN bus sample frequency is 1 Hz, the vision-based system operates in real-time at 25 frames per second (25 Hz). The proposed data fusion scheme provides information at the lowest sample frequency (1 Hz) covering two consecutive GPS measurements, the vehicle speed v_i^h (via CAN bus) and the outputs of the vision module.

The outputs of the side, forward, and rear vehicle detection systems at frame i are the number of detected vehicles N_i and their corresponding distances to the host vehicle $d_i^{(k)}$ (note that d is used here as a distance/range

FIGURE 22: GPS trajectory and the corresponding traffic load computed at the central unit (the aerial image has been obtained from Google Earth).

measurement). These outputs are combined to cover the whole local environment of the vehicle. The traffic load at frame i is given by next expression

$$L_i = \frac{N_i + 1}{N_{\text{MAX}}}, \tag{7}$$

where N_{MAX} is the maximum number of vehicles in range that can be detected by the three systems (in our case N_{MAX} is defined as 9 or 13 for two lanes and three lanes roads, resp.). The average road speed at frame i is computed as follows:

$$
\begin{aligned}
v_i &= \frac{1}{N_i + 1} \left(\sum_{k=0}^{N_i - 1} \left(\frac{\left(d_i^{(k)} - d_{i-1}^{(k)} \right)}{\Delta t} + v_i^h \right) + v_i^h \right) \\
&= \frac{1}{(N_i + 1)\Delta t} \sum_{k=0}^{N_i - 1} \left(d_i^{(k)} - d_{i-1}^{(k)} \right) + v_i^h,
\end{aligned}
\tag{8}
$$

where $d_i^{(k)}$ and $d_{i-1}^{(k)}$ represent the distance between the host vehicle and vehicle k at frames i and $i - 1$, respectively, Δt corresponds to the sample time, v_i^h is the host vehicle speed provided by the CAN bus, and N_i is the number of detected vehicles. Note that the distance values correspond to filtered measurements since they are obtained from the first two elements of the Kalman filter state vector (u and v) using known camera geometry and ground-plane constraints.

Two consecutive GPS measurements define both a spatial and a temporal segment. The temporal segment corresponds to the GPS sample time (1 second), and the spatial segment will be defined as the globally referenced trajectory between the two GPS measurements. In order to obtain the extended FCD information (i.e., the road traffic load and the road speed) for this spatio/temporal segment we integrate the values supplied by the vision modules during 25 consecutive frames. With this approach a dense coverage of the road traffic load and the road speed can be assured for host vehicle speeds up to 180 km/h since the total range of the vision

module covers more than 50 m (25 meters for both the rear and the forward looking modules; the side range covers up to two third parts of the bus length in the adjacent lane). Obviously this maximum speed will never be exceeded by a public bus. This approach facilitates further map-matching tasks since the extended FCD information between two consecutive points will always be globally referenced.

4. Experimental Results

The system was implemented on a PC Core 2 Duo at 3.0 GHz and tested in real traffic conditions using CMOS cameras with low-resolution images (320×240). After training and test, a tradeoff point has been chosen at detection rate (DR) of 95% and false positive rate (FPR) of 5% for the rear-SVM classifier and at DR of 90% and FPR of 6% for the forward-SVM classifier. We have to note that these numbers are obtained in an offline single-frame fashion, so that, they will be improved in subsequently stages. In addition, the lane detection system reduces the searching area and the number of false candidates passed to further stages.

In order to validate the proposed vision-based vehicle detection system as an extended source for FCD applications we have recorded several video sequences in real traffic conditions, and we have manually labeled the number of vehicles in range at every frame (a total of 800 frames). The speed of the host vehicle was around 90 km/h so the length of the traveled route was 1 km approximately. Both the traffic load L_i and the average road speed v_i are computed at every frame using (7) and (8). Figure 16 shows the ground truth and the number of vehicles detected in range. Most of the errors take places in cases where the host vehicle is passing beneath a bridge due to strong illumination changes (see Figure 17) and in curves or cases where there are strong changes in the vehicle pitch, roll or camera height.

The traffic load L_i and the average road speed v_i at every frame are depicted in Figures 18 and 19, respectively.

These values are provided by the vision modules at a frequency of 25 Hz. As the extended FCD information is supplied to the central unit at a frequency of 1 Hz the traffic load and the average road speed are finally integrated during 25 consecutive frames. These results are shown in Figures 20 and 21. We use a colour code to describe the level of traffic load and the road speed: *green* indicates that there is good flow/high speeds; *yellow* indicates that there is semi-dense traffic/medium speeds, and *red* shows dense traffic/slow speeds (traffic jams). After combining the results with the GPS measurements we can obtain the traffic load in universal transverse mercator (UTM) coordinates, as depicted in Figure 22 (note that map-matching is not carried out; the aerial image has been obtained from Google Earth).

5. Concluding Remarks

This paper presented a complete vision-based vehicle detection system that enhances the data supplied by FCD systems in the context of vehicular ad hoc networks. The system is composed of three vision subsystems (side, forward and rear subsystems) that detect the traffic load and the relative velocities of the vehicles contained in the local area of the host vehicle. Under certain constraints, such as good weather and daytime conditions, absolute velocities, and global positioning are obtained after combining the outputs provided by the vision modules with the outputs supplied by the CAN Bus and the GPS sensor. Standard FCD systems provide the vehicle position, speed, and direction. The proposed approach extends this information by including more representative measurements corresponding to the traffic load and the average road speed.

In order to cover the entire road network, the proposed vision-based system is defined for being installed onboard a fleet of public buses where privacy is a minor issue. The extended packets collected by each moving vehicle are transmitted to the central unit by means of a GPRS/UMTS data connection. The central unit merges the extended FCD in order to maintain an updated map of the traffic conditions (traffic load and average road speed).

The presented experiments are promising in terms of detection performance and computational costs. However, significant effort is further necessary before deploying a system for large-scale real applications. For this purpose, new experiments will be carried out merging the data collected by more than one vehicle, including map-matching techniques and further analysis on V2I and V2V communications (e.g., using repetition based MAC protocols [13]). In addition, the proposed vision-based vehicle detection system will be extended to deal with complex weather conditions (e.g., wet or snowy roads) as well as night-time conditions.

Acknowledgments

References

[1] R. Bishop, *Intelligent Vehicle Technologies and Trends*, Artech House, Boston, Mass, USA, 2005.

[2] S. Messelodi, C. M. Modena, M. Zanin et al., "Intelligent extended floating car data collection," *Expert Systems with Applications*, vol. 36, no. 3, part 1, pp. 4213–4227, 2009.

[3] S. Rass, S. Fuchs, M. Schaffer, and K. Kyamakya, "How to protect privacy in floating car data systems," in *Proceedings of the 5th ACM International Workshop on Vehicular Inter-Networking (VANET '08)*, pp. 17–22, September 2008.

[4] P. Day, J. Wu, and N. Poulton, *Briefing Note on Floating Car Data*, ITS Internacional, 2006.

[5] M. A. Sotelo, J. Nuevo, L. M. Bergasa, M. Ocaña, I. Parra, and D. Fernández, "Road vehicle recognition in monocular images," in *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE '05)*, pp. 1471–1476, June 2005.

[6] M. Á. Sotelo and J. Barriga, "Blind spot detection using vision for automotive applications," *Journal of Zhejiang University: Science A*, vol. 9, no. 10, pp. 1369–1372, 2008.

[7] C. J. Christopher Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 886–893, June 2005.

[9] Y. Derong, Z. Yuanyuan, and L. Dongguo, "Fast computation of multiscale morphological operations for local contrast enhancement," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE-EMBS '05)*, pp. 3090–3092, September 2005.

[10] E. R. Dougherty, *An Introduction to Morphological Image Processing*, SPIE Optical Engineering Press, 1992.

[11] D. Balcones, D. F. Llorca, M. A. Sotelo, et al., "Real-time vision-based vehicle detection for rear-end collision mitigation systems," in *Proceedings of the 12th International Conference on Computer Aided Systems Theory (EUROCAST '09)*, vol. 5717 of *Lecture Notes in Computer Science*, pp. 320–325, 2009.

[12] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering, Series D*, vol. 82, pp. 35–45, 1960.

[13] B. Hassanabadi, L. Zhang, and S. Valaee, "Index coded repetition-based MAC in vehicular ad-hoc networks," in *Proceedings of the 6th IEEE Conference on Consumer Communications and Networking Conference*, 2009.

EUSIPCO 2011

19th European Signal Processing Conference
August 29- September 2, 2011, Barcelona (Spain)

Photograph © Turisme de Barcelona / J. Trullàs

Preliminary call for papers

The 2011 European Signal Processing Conference (EUSIPCO-2011) is the nineteenth in a series of conferences promoted by the European Association for Signal Processing (EURASIP, www.eurasip.org). This year edition will take place in Barcelona, capital city of Catalonia (Spain), and will be jointly organized by the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) and the Universitat Politècnica de Catalunya (UPC).

EUSIPCO-2011 will focus on key aspects of signal processing theory and applications as listed below. Acceptance of submissions will be based on quality, relevance and originality. Accepted papers will be published in the EUSIPCO proceedings and presented during the conference. Paper submissions, proposals for tutorials and proposals for special sessions are invited in, but not limited to, the following areas of interest.

Areas of Interest

- Audio and electro-acoustics.
- Design, implementation, and applications of signal processing systems.
- Multimedia signal processing and coding.
- Image and multidimensional signal processing.
- Signal detection and estimation.
- Sensor array and multi-channel signal processing.
- Sensor fusion in networked systems.
- Signal processing for communications.
- Medical imaging and image analysis.
- Non-stationary, non-linear and non-Gaussian signal processing.

Submissions

Procedures to submit a paper and proposals for special sessions and tutorials will be detailed at www.eusipco2011.org. Submitted papers must be camera-ready, no more than 5 pages long, and conforming to the standard specified on the EUSIPCO 2011 web site. First authors who are registered students can participate in the best student paper competition.

Important Deadlines:

Proposals for special sessions	15 Dec 2010
Proposals for tutorials	18 Feb 2011
Electronic submission of full papers	**21 Feb 2011**
Notification of acceptance	23 May 2011
Submission of camera-ready papers	6 Jun 2011

Webpage: www.eusipco2011.org

Organizing Committee

Honorary Chair
Miguel A. Lagunas (CTTC)

General Chair
Ana I. Pérez-Neira (UPC)

General Vice-Chair
Carles Antón-Haro (CTTC)

Technical Program Chair
Xavier Mestre (CTTC)

Technical Program Co-Chairs
Javier Hernando (UPC)
Montserrat Pardàs (UPC)

Plenary Talks
Ferran Marqués (UPC)
Yonina Eldar (Technion)

Special Sessions
Ignacio Santamaría (Unversidad de Cantabria)
Mats Bengtsson (KTH)

Finances
Montserrat Nájar (UPC)

Tutorials
Daniel P. Palomar
(Hong Kong UST)
Beatrice Pesquet-Popescu (ENST)

Publicity
Stephan Pfletschinger (CTTC)
Mònica Navarro (CTTC)

Publications
Antonio Pascual (UPC)
Carles Fernández (CTTC)

Industrial Liaison & Exhibits
Angeliki Alexiou
(University of Piraeus)
Albert Sitjà (CTTC)

International Liaison
Ju Liu (Shandong University-China)
Jinhong Yuan (UNSW-Australia)
Tamas Sziranyi (SZTAKI -Hungary)
Rich Stern (CMU-USA)
Ricardo L. de Queiroz (UNB-Brazil)