2013 IEEE Intelligent Vehicles Symposium (IV)
June 23-26, 2013, Gold Coast, Australia

Vision-based parking assistance system for leaving perpendicular and angle parking lots

D. F. Llorca, S. Álvarez, M. A. Sotelo

***Abstract*— Backing-out maneuvers in perpendicular or angle parking lots are one of the most dangerous maneuvers, specially in cases where side parked cars block the driver view of the potential traffic flow. In this paper a new vision-based Advanced Driver Assistance System (ADAS) is proposed to automatically warn the driver in such scenarios. A monocular gray-scale camera is installed at the back-right side of the vehicle. A Finite State Machine (FSM) defined according to three CAN-Bus variables and a manual signal provided by the user is used to handle the activation/deactivation of the detection module. The proposed oncoming traffic detection module computes spatio-temporal images from a set of pre-defined scan-lines which are related to the position of the road. A novel spatio-temporal motion descriptor is proposed (STHOL) accounting the number of lines, their orientation and length of the spatio-temporal images. A Bayesian framework is used to trigger the warning signal using multivariate normal density functions. Experiments are conducted on image data captured from a vehicle parked at different locations of an urban environment, including different lighting conditions. We demonstrate that the proposed approach provides robust results maintaining processing rates close to real-time.**

***Index Terms*— Park Assist, Perpendicular and Angle Parkings, Backing-out Maneuvers, Spatio-temporal Images, Motion Patterns, ADAS.**

I. INTRODUCTION

In the last years, a considerable number of research works and industrial developments on Intelligent Parking Assist Systems (IPAS) have been proposed, including both assistance and automatic parking approaches. Most of these systems have been designed to assist the driver when parking in parallel, perpendicular or angle parking lots. However the development of intelligent systems designed to assist the driver when leaving the parking lots has been somewhat neglected in the literature.

The nature of parking assistance systems for entering a parking lot is different from that of parking assistance systems for backing-out manoeuvres. On the one hand, the main goal of IPAS that assist drivers when parking is to ease the maneuver avoiding small collisions, reducing car damage, and avoiding personal injuries. Although the number of injured people is not negligible at all (more than 6.000 people are injured yearly by vehicles that are backing up only in the United States [1]), the low speed of the vehicles involved in the accidents reduce the severity of the damage. On the other hand, leaving parking manoeuvres imply to enter in an active traffic lane where vehicles move at a relative speed much higher than the speed of the vehicle that is leaving the parking lot. This situation can be particularly dangerous when the pull out manoeuvre has to be done blindly, since the driver does not have visibility of the oncoming traffic. In other words, the safety component of IPAS devised to assist the driver when leaving a parking space is much more relevant since the possible collisions may cause serious injuries and damages.

In this paper a new vision-based Advanced Driver Assistance System (ADAS) is proposed to deal with scenarios like the ones depicted in Figs. 1(a) and 1(b). We consider backing-out or heading-out maneuvers in perpendicular or angle parking lots, in cases where side parked cars block the driver view of the potential traffic flow. In such cases the common recommendation can be simplified as to move slow looking at every direction, but it is not possible to avoid initiating the maneuver in blind conditions. We propose a vision-based solution using a camera located at the back-right side of the vehicle [1] which captures images with a better Field of View (FOV) than the driver's FOV (see Fig. 1).

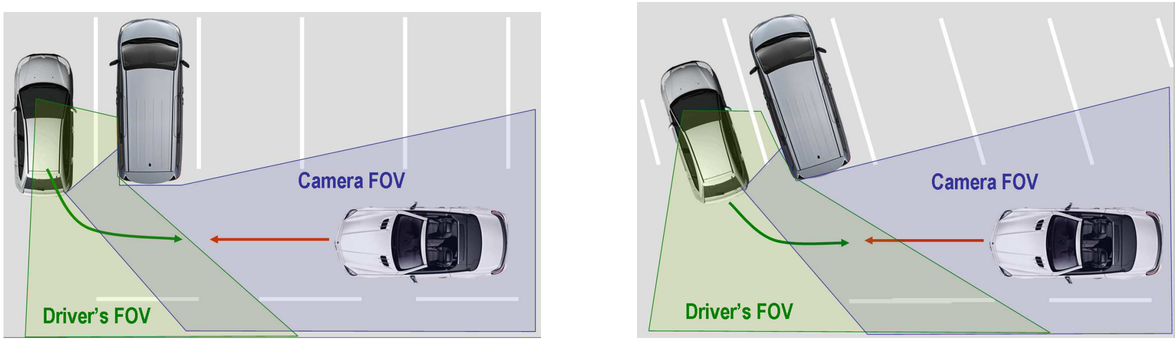
II. RELATED WORK

A sizeable body of literature exists related to IPAS, including range sensor-based approaches [2], monocular-based systems [3], [4], and motion stereo-based proposals [5]. However all these systems propose target position-designation methods to assist the driver when parking or to perform automatic parking. The closest field related with our proposal can be found in the area of Blind Spot Detection systems (BSD) that monitor the road behind and next to the host vehicle, warning the driver when there are vehicles in the blind spot of the side-view. These systems are mainly based on the use of cameras installed in the left and/or right door mirrors [6]. These systems can be utilized to assist the driver when leaving a parallel parking lot, but the position of the camera makes not possible to use BSD systems in the scenarios depicted in Fig. 1. In addition, BSD systems usually take advantage of the opposite direction between the implicit optical flow and the motion of the overtaking vehicles. This difference is not so evident in the scenarios used in this work.

Considering the recognition of vehicles in the context of ADAS, extensive literature is available for both forward and rear vehicle detection [7]. The FOV of the camera and the

D. F. Llorca, S. Álvarez, M. A. Sotelo are with the Computer Engineering Department, Polytechnic School, University of Alcalá, Madrid, Spain. email: llorca, sergio.alvarez, sotelo@aut.uah.es.

[1] In countries with left-hand traffic the camera will be located at the back-left side of the vehicle



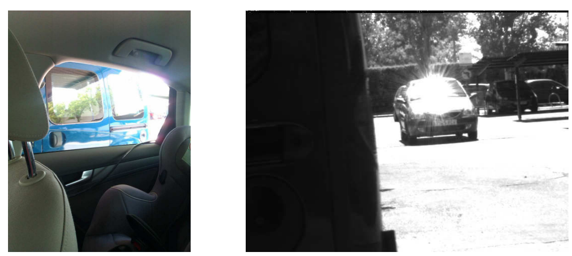
(a) Back-out perpendicular parking (b) Back-out angle parking

Fig. 1. Driver and camera Field of View (FOV).

type of maneuver when leaving a perpendicular or angle parking (see Fig. 1) provide images similar to the ones used by rear vehicle detection systems [8]. Most of these systems follow a three-staged framework: Region-Of-Interest (ROI) generation (monocular or stereo [9])), classification and tracking. All these stages are needed since the system has to deal with a wide number of scenarios and driving conditions. However, in the context of our application, the number of possible scenarios is much lower so we aim to devise a simpler system without this tree-staged scheme.

We propose a probabilistic model of the spatio-temporal motion patterns obtained from a set of virtual lines placed following the road location. The spatio-temporal domain is analyzed by accounting the number of lines and their length with respect to their orientations in a histogram of orientations that we so-called Spatio-Temporal Histograms of Oriented Lines (STHOL). The resulting feature vectors are modeled assuming a normalized multivariate Gaussian distribution for two types of scenarios (classes): *oncoming traffic* and *free road*. Bayes decision theory is then used by means of discriminant functions based on the minimum error rate that assumes equal prior probabilities. Finally, if the p.d.f. of the *oncoming traffic* class is larger than *free traffic* class p.d.f, the system triggers a warning signal that alerts the driver of oncoming traffic.

Fig. 2. Camera located at the back-right side of the vehicle.



(a) Driver's point of view (b) Image captured by the camera

Fig. 3. Driver and camera point of view.

III. SYSTEM DESCRIPTION

The proposed architecture of the system is composed of three main parts: camera, processor and CAN-Bus communications. A gray-scale 640×480 resolution camera is used, with a focal lenght of 12.5mm. The location of the camera is depicted in Fig. 2. This is obviously a preliminary structure since the camera should be integrated inside the vehicle bodywork. As can be observed in Figs. 3(a) and 3(b) the point of view of the camera is much better than the driver's point of view.

The processor is a PC-based architecture that is connected with both the camera and the CAN-Bus interface. From the CAN-Bus we obtain the next variables: *steering angle*, *car speed* and *current gear*. These variables are used to trigger on/off the detection module according to the Finite State Machine (FSM) described in Fig. 4. As can be observed the system has to be firstly activated by the user. Then the system waits until the car has been put into reverse gear and the detection module is then triggered on. The system stops if one of the following conditions are met: (1) vehicle speed is greater than 5km/h or (2) steering angle is greater than 10 degrees with respect to the zero reference position or (3) reverse gear is deactivated.

IV. SPATIO-TEMPORAL DETECTION MODEL

An overview of the proposed spatio-temporal detection model of the oncoming traffic is depicted in Fig. 5. Spatio-temporal images are computed using a pre-defined grid of scan-lines which are related with the location of the road. These images are then analyzed using a line detection stage, which provides the lines, their orientation and length. This information is used to compute the so-called Spatio-

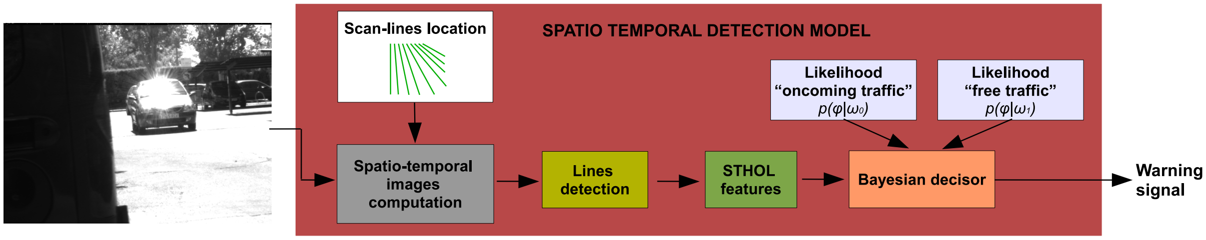


Fig. 5. Overview of the spatio-temporal detection module.

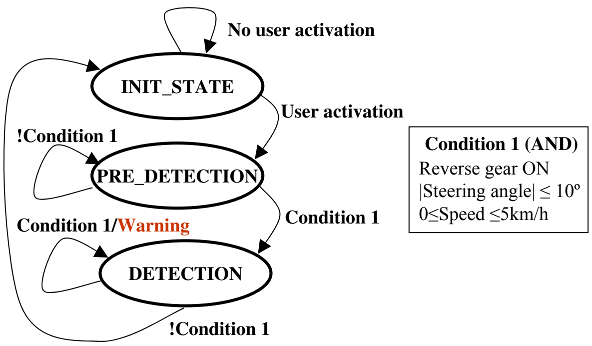


Fig. 4. FSM for detection module.

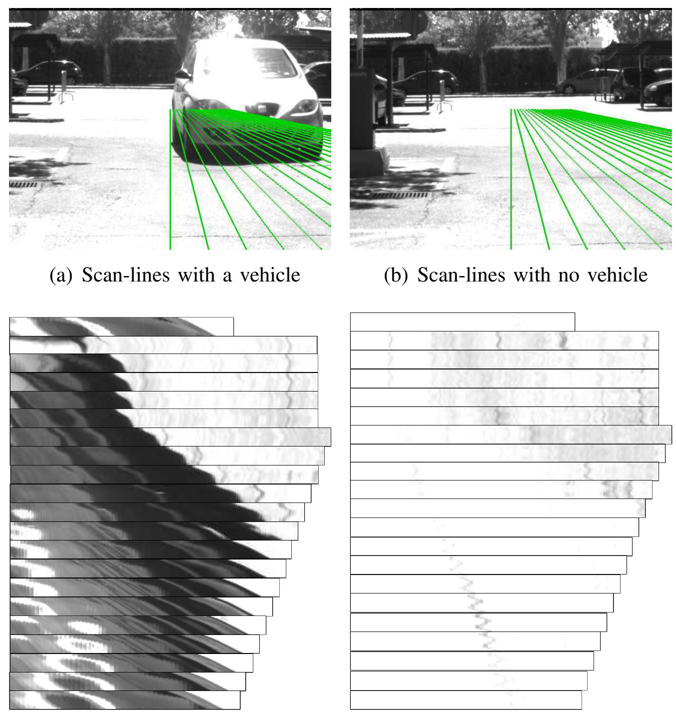
Temporal Histograms of Oriented Lines (STHOL), which are the features used to represent the current state of the adjacent lane: oncoming traffic or free traffic. Finally, a Bayesian decision scheme is used to trigger the warning signal to the user. In the following, details of each one of the modules represented in Fig. 5 are given.

A. Spatio-temporal images

Vehicle detection proceeds with the computation of spatio-temporal images which represents a single intensity scan-line collected over several frames. This approach was presented in [10] to perform crow detection in video sequences using a set of horizontal scan-lines. In our case, the distribution of the scan-lines follows a pre-defined representation of the road using the flat world assumption, extrinsic parameters of the camera w.r.t. the road (obtained by means of an off-line camera calibration process) and a pre-defined grid which covers half of the road. The definition of the number of scan-lines and their distribution have been experimentally determined taking into account the maximum and minimum range, as well as a trade-off between computation time and the density of information. Two examples are depicted in Figs. 6(a) and 6(b) where we can observe that only the half of the image is covered [2].

For each scan-line we create a spatio-temporal image that contains that scan-line in the last 16 frames (the scan-line from the last image is placed at the upper part of the spatio-temporal image and the rest of the scan-lines are shifted to the bottom). As can be observed in Figs. 6(c) and 6(d) the motion patterns showed by the spatio temporal images between the case of a vehicle approaching and no vehicle approaching are, at first glance, very different.

[2]In countries with left-hand traffic the definition of the scan lines will be symmetric and located at the other side.



(a) Scan-lines with a vehicle
(b) Scan-lines with no vehicle
(c) Spatio-temporal images with a vehicle
(d) Spatio-temporal images with no vehicle

Fig. 6. Two examples of the scan-lines and spatio temporal images. Note that the size of the spatio-temporal images is different depending on the scan-line.

B. Feature selection

Given a set of spatio-temporal images, a new descriptor is here introduced by accounting the number of lines and their length with respect to their orientations in a histogram of orientations that we denote as Spatio-Temporal Histograms of Oriented Lines (STHOL). Instead of using the Hough transform as in [10], which in our case provides noisy results, we propose to use the approach suggested by [11]. The first step of the line detection is the computation of the image derivatives using Sobel edge detector. The gradient direction is then quantized into a set of k ranges (in our

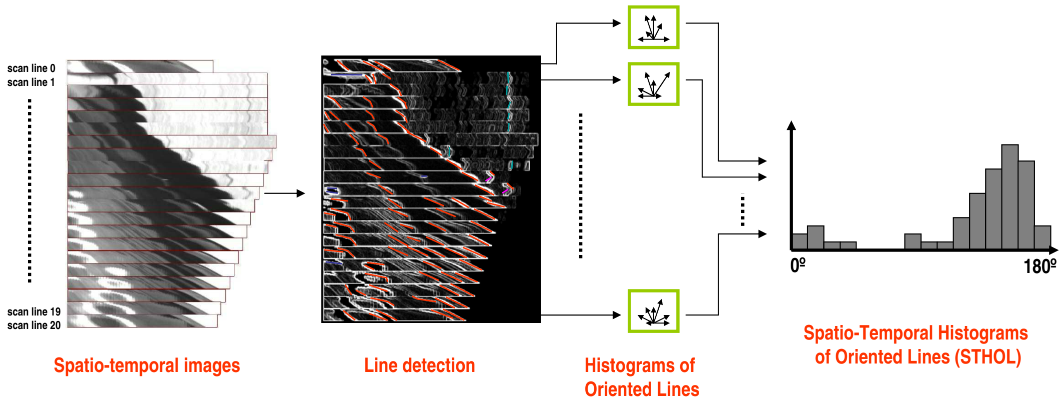


Fig. 7. Overview of the STHOL feature selection architecture.

case, $k = 16$) where all the edge pixels having an orientation within the specific range fall into the corresponding bin and will be properly labeled. The edge pixels having the same label are then grouped together using connected components algorithm. The line segment candidates are obtained by fitting a line parameterized by an angle θ and a distance from the origin ρ using the following expression:

$$\rho = x\cos\theta + y\sin\theta \tag{1}$$

Each obtained connected component is a list of edge pixels (u_i, v_i) with similar gradient orientation, which is considered as the line support regions. The line parameters are then determined from the eigenvalues λ_1 and λ_2 and eigenvectors $\vec{v}_1$ and $\vec{v}_2$ of the matrix D associated with the line support region which is given by:

$$D = \begin{bmatrix} \sum_i (x_i-\bar{x})_i^2 & \sum_i (x_i-\bar{x})(y_i-\bar{y}) \\ \sum_i (x_i-\bar{x})(y_i-\bar{y}) & \sum_i (y_i-\bar{y})^2 \end{bmatrix} \tag{2}$$

where $\bar{x} = \frac{1}{n}\sum_i x_i$ and $\bar{y} = \frac{1}{n}\sum_i y_i$ are the mid-points of the line segment. The second eigenvalue of an ideal line should be zero. The quality of the lines fit is modeled by the ratio of the two eigenvalues of matrix D, i.e., $\frac{\lambda_1}{\lambda_2}$. If the eigenvector $\vec{v}_1$ is associated with the largest eigenvalue, the line parameters (ρ, θ) are determined using:

$$\begin{aligned} \theta &= atan2\left(\vec{v}_1(2), \vec{v}_1(1)\right) \\ \rho &= \bar{x}\cos\theta + \bar{y}\sin\theta \end{aligned} \tag{3}$$

This procedure is applied on each one of the spatio-temporal images, providing a set of lines with their orientation and length. Motion patterns corresponding to oncoming traffic yield a considerable number of lines with a specific orientation that clearly differs from cases without oncoming vehicles (see Figs. 6(c) and 6(d)). The number of lines detected on each spatio-temporal image is then combined in an orientation histogram with d bins evenly spaced over 0°-180° (unsigned gradient, i.e., the sign of the line is ignored). To take into account the strength of each line, votes are directly related with the length of the line. Thus each image, which integrates information from the last 16 frames, provides a specific d-dimensional feature vector that accounts for the number of lines, their lengths and their orientation corresponding to the spatio-temporal images of all the pre-defined scan-lines. We have called this feature vector Spatio-Temporal Histograms of Oriented Lines (STHOL). An overview of the proposed architecture is depicted in Fig. 7.

C. Bayesian decision scheme

Given a particular image I that contains temporal information of the last 15 frames, our aim is to estimate its posterior probability, $P(\omega_0|I)$ with respect to the *oncoming traffic* class ω_0. To that extend, we represent the image I in terms of STHOL features φ_I and follow a Bayesian approach considering the *free traffic* class:

$$P(\omega_0|I) = P(\omega_0|\varphi_I) = \frac{p(\varphi_I|\omega_0)P(\omega_0)}{\sum_{i=0}^{1} p(\varphi_I|\omega_i)P(\omega_i)} \tag{4}$$

Although, it may be intuitive to consider the *free traffic* class to be more probable than the *oncoming traffic* class, priors for both *oncoming traffic* ω_0 and *free traffic* ω_1 class, $P(\omega_0)$ and $P(\omega_1)$ are considered uniform and equal. This is an obvious simplification that could be overcome by, for example, modeling the priors using traffic data depending on the time and the global positioning where the vehicle is parked. We consider this interesting analysis out of the scope of this paper. Accordingly, since we manage equal priors, and the evidence $\sum_{i=0}^{1} p(\varphi_I|\omega_i)P(\omega_i)$ is also common to both classes, our problem can be simplified by estimating and evaluating the likelihoods $p(\varphi_I|\omega_i)$ which represent the probability of a particular observation (feature descriptor) given the traffic state of the lane (*oncoming traffic* or *free traffic*).

The following multivariate normal density function is used to model the likelihoods, $p(\varphi_I|\omega_i) \sim N(\mu_i, \Sigma_i)$:

$$p(\varphi_I|\omega_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} exp\left[-\frac{1}{2}(x-\mu_i)^t\Sigma_i^{-1}(x-\mu_i)\right] \tag{5}$$

where x is the d-component feature descriptor (STHOL), μ_i is the d-component mean vector for class ω_i and Σ_i is the d-by-d covariance matrix corresponding to class ω_i. The next parameters are then estimated using the training data: sample means μ_0 and μ_1 and sample covariance matrices Σ_0 and Σ_1.

We finally use the minimum-error-rate classification using the discriminant function:

$$g_i(x) = lnp(\varphi_I|\omega_i) + lnP(\omega_i) \tag{6}$$

By merging Eq. 5 and Eq. 6 we have:

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^t\Sigma_i^{-1}(x-\mu_i) - \frac{d}{2}ln2\pi - \frac{1}{2}ln|\Sigma_i| + lnP(\omega_i) \tag{7}$$

Taking into account that we consider equal priors, and equal feature vector dimension for each class, the terms $(d/2)ln2\pi$ and $lnP(\omega_i)$ can be dropped from Eq. 7, giving the following discriminant function:

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^t\Sigma_i^{-1}(x-\mu_i) - \frac{1}{2}ln|\Sigma_i| \tag{8}$$

Instead of using two discriminant functions g_0 and g_1, and assigning φ_I to ω_0 if $g_0 > g_1$ we define a single discriminant function $g(x) = g_0(x) - g_1(x)$ and we finally trigger the warning signal if $g(x) > \theta_{TH}$.

V. EXPERIMENTS

The proposed oncoming vehicle detection approach to assist the driver when leaving a perpendicular or angle parking was tested in experiments with data recorded from a real vehicle in real urban traffic conditions. Four different locations have been used, including different levels of visibility due to the size of the side parked vehicles and different camera orientations. Datasets were acquired in daylight conditions. Some examples of the different locations and lighting conditions contained in our dataset are depicted in Fig. 8.

Fig. 8. Sample images of the datasets.

The experimental data is divided in two datasets. One of the datasets is utilized at a time to learn the probabilistic spatio-temporal model. Performance is evaluated in the remaining dataset. To evaluate the quality of the proposed method, we have labeled all the images in two categories: oncoming traffic and free traffic. Note that vehicles that are out of the range of the vision system (50m with our configuration) were labeled as free traffic until they enter in the range of the camera. Table I depicts the number of images of the two datasets, including the number of images with free traffic conditions, the number of images with oncoming traffic as well as the number of vehicle trajectories (one vehicle usually appears a number of frames which is directly related with its speed). In addition, stationary cars that appear inside the range area, are considered as free traffic. The proposed method should be able to distinguish this specific case.

TABLE I

STATISTICS OF THE CONSIDERED DATA SETS.

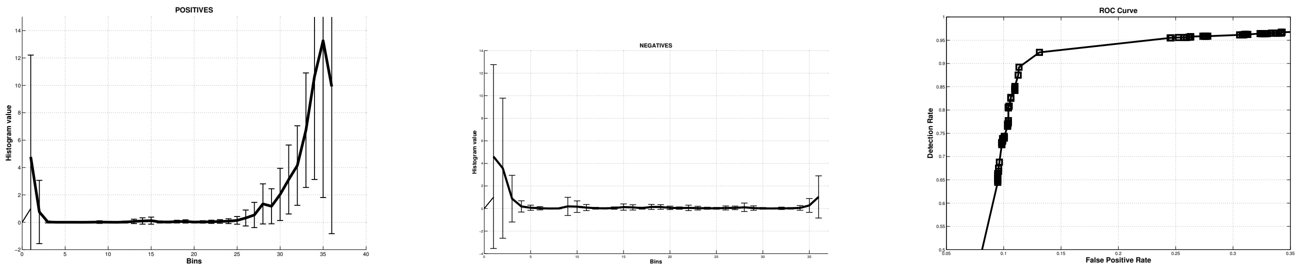
	dataset 1	dataset 2
# of images	16124	6902
# of free traffic images	12862	4838
# of oncoming traffic images	3262	2064
# of vehicle trajectories	34	15

In our case the number of bins used in the STHOL features has been experimentally fixed to 36. The mean values of the multivariate normal density function as well as their standard deviations (computed as the squared root of the diagonal elements of the covariance matrices) for both oncoming traffic and free traffic classes are depicted in Figs. 9(a) and 9(b). As can be observed, the mean values of the multivariate Gaussian modeling corresponding to the STHOL features are very different for both classes. The orientations of most of the lines when a vehicle is approaching lie between $120° - 170°$. Fig. 9(c) shows the performance of the proposed classifier in terms of ROC curve, by varying the threshold value θ_{TH} of the discriminant function. Note that these results correspond to single-frame classification. A deeper analysis of the errors show that most of the false negatives and false positives appear when the vehicle is located far from the vehicle around the limits of the system range, obtaining a more robust detection as the vehicle gets closer to the camera. In such cases the number of usable edges and the length of the lines are insufficient to take a consistent decision. In practice the system correctly triggers the warning signal for all the vehicles that appear in our test data set (15 vehicles trajectories). Fig. 10 depicts some results of the proposed system. Processing time is around 20Hz using a C/C++ implementation on a state-of-the-art 2.66 GHz Intel PC.

VI. CONCLUSION

This paper presented a novel solution to a new type of ADAS to automatically warn the driver when backing-out in perpendicular or angle parking lots, specially in cases where side parked cars block the driver view of the potential traffic flow. The detection system is handled by a FSM. A novel spatio-temporal motion descriptor is presented (STHOL features) to robustly represent oncoming traffic or free traffic states. A Bayesian framework is finally used to trigger the warning signal.

Future work will be concerned with the evaluation of the method in night-time conditions, comparisons between our generative approach and other discriminative approaches such as SVM-based. A more sophisticated approach should be elaborated to model the priors using massive traffic data globally and temporally referenced, since it is obvious that the prior probability of meeting oncoming traffic depends on variables such as the time of the day, the type of road,



(a) (b) (c)

Fig. 9. Mean value of the multivariate Gaussian and standard deviations (square root of the diagonal elements of the covariance matrix) corresponding to (a) Oncoming traffic and (b) Free traffic classes. (c) Receiver-Operating-Characteristic curve of the single frame Bayesian classifier.

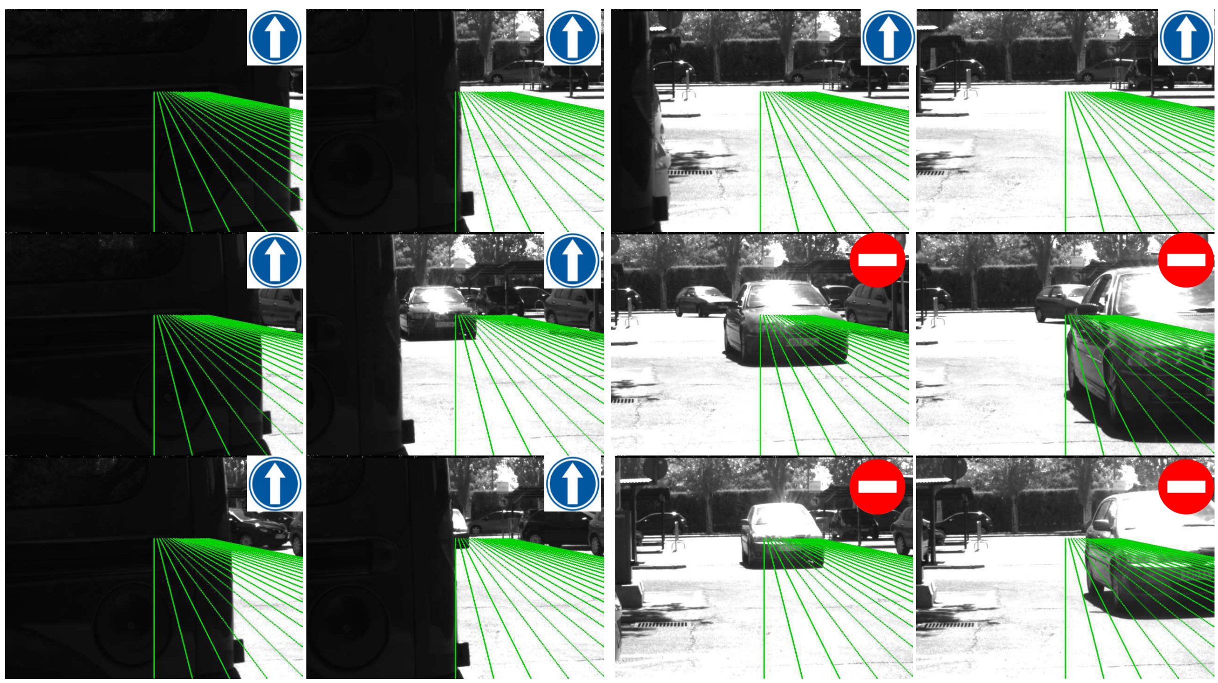


Fig. 10. Three sample sequences with the triggered warning signal.

etc. More experimental work should be carried out including optimization procedures and different configurations of the STHOL feature descriptor.

VII. ACKNOWLEDGMENTS

This work was supported by the Spanish Ministry of Science and Innovation under Research Grant ONDA-FP TRA2011-27712-C02-02.

REFERENCES

[1] NHTSA, "Vehicle backover avoidance technology," National Highway Traffic Safety Administration, Madrid, Spain, Report to Congress, 2006.

[2] H. G. Jung, Y. H. Cho, P. J. Yoon, and J. Kim, "Scanning laser radarbased target position designation for parking aid system," *IEEE Trans. on Intell. Transp. Sys.*, vol. 9, pp. 406–424, 2008.

[3] H. G. Jung, D. S. Kim, P. J. Yoon, and J. Kim, "Parking slot markings recognition for automatic parking assist system," in *IEEE Intelligent Vehicle Symposium*, 2006.

[4] H. G. Jung, D. S. Kim, and J. Kim, "Light-stripe-projection-based target position designation for intelligent parking-assist system," *IEEE Trans. on Intell. Transp. Sys.*, vol. 11, pp. 942–953, 2010.

[5] ——, "Automatic free parking space detection by using motion stereo-based 3d reconstruction," *Machine Vision and Applications*, vol. 21, pp. 163–176, 2010.

[6] M. A. Sotelo, J. Barriga, D. Fernández, I. Parra, J. E. Naranjo, M. Marrón, S. Álvarez, and M. Gavilán, "Vision-based blind spot detection using optical flow," *EUROCAST 2007, LNCS*, vol. 4739, pp. 1113–1118, 2007.

[7] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection: A review," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 694–711, 2006.

[8] D. Balcones, D. F. Llorca, M. A. Sotelo, M. Gavilán, S. Álvarez, I. Parra, and M. Ocana, "Real-time vision-based vehicle detection for rear-end collision mitigation systems," *EUROCAST 2009, LNCS*, vol. 5717, pp. 320–235, 2009.

[9] D. F. Llorca, M. A. Sotelo, A. M. Hellín, A. Orellana, M. Gavilán, I. G. Daza, and A. G. Lorente, "Stereo regions-of-interest selection for pedestrian protection: A survey," *Transportation Research Part C*, vol. 25, pp. 226–237, 2012.

[10] P. Reisman, O. Mano, S. Avidan, and A. Shashua, "Crowd detection in video sequences," in *IEEE Intelligent Vehicle Symposium*, 2004.

[11] W. Zhang and J. Kosecka, "Efficient computation of vanishing points," in *IEEE International Conference on Robotics and Automation*, 2002.