Smart Side View Mirror Camera for Real Time System

Nunziata Ivana Guarneri, Arcangelo Bruna, Giuseppe Spampinato, Antonio Buemi

Abstract—In the last decade, automotive companies have invested a lot in terms of innovation about many aspects regarding the automatic driver assistance systems. One innovation regards the usage of a smart camera placed on the car's side mirror for monitoring the back and lateral road situation. A common road scenario is the overtaking of the preceding car and, in this case, a brief distraction or a loss of concentration can lead the driver to undertake this action, even if there is an already overtaking vehicle, leading to serious accidents. A valid support for a secure drive can be a smart camera system, which is able to automatically analyze the road scenario and consequentially to warn the driver when another vehicle is overtaking. This paper describes a method for monitoring the side view of a vehicle by using camera optical flow motion vectors. The proposed solution detects the presence of incoming vehicles, assesses their distance from the host car, and warns the driver through different levels of alert according to the estimated distance. Due to the low complexity and computational cost, the proposed system ensures real time performances.

Keywords—Camera calibration, ego motion, kalman filters, object tracking, real time systems.

I. INTRODUCTION

DUE to the increasing attention toward the automotive smart systems, many efforts have been spent in terms of new hardware and software car's equipment. Today, the newest cars present in the market make use of back, forward or side cameras for different purposes. Some of the most popular applications are: the Cross Traffic Alert (CTA) [1], the Lane Departure Warning (LDW) [2], [3], the Collision Avoidance (CA) [4]-[6], and the Blind Spot Detection (BSD) [7], [8]. The different Advanced Driver Assistance systems (ADAS) topics make sense in different road scenarios, for example, the CTA is highly useful in city road environment where other vehicles can cross the road, conversely the LDW or the BSD can be highly useful in highways where the car reaches high speeds and a brief distraction of the driver can lead to an accident. The technologies used in ADAS systems are different. The car's camera based systems are employed in many fields as the surveillance of the vehicle [9], in robotic/navigation [10], for the traffic signs recognition [11] and pedestrians tracking [12]. The radar based systems allow accurate and robust measurement of the distance between vehicles to avoid collision [13]. The far infrared based systems are able to detect living objects like animals and pedestrians [14]. These sensors are often used in different numbers and

I. Guarneri, A. Bruna, G. Spampinato and A. Buemi are with the STMicroelectronics, Catania, Italy (e-mail: ivana.guarneri@st.com, arcangelo.bruna@st.com, giuseppe.spampinato@st.com, antonio.buemi@st.com).

combinations to realize really performing systems. In this paper, a system is presented for monitoring the lateral view of the vehicle (BSD use case) and alerting the driver in case of a possible risk. It is able to detect a vehicle in overtaking phase, for assessing its distance from the host car, and according to the estimated distance, to warn the driver. This is a camera based system and it uses exclusively the motion vectors extracted from the scene. The reason to use only the optical flow is due to the fact that, today, a growing number of manufactures produce image sensors with hardware implemented optical flow. It allows obtaining real time applications with low extra computational effort. Moreover, by exploiting processors directly embedded in the same chip, the system can work directly in the Image Signal Processor to avoid overloading the ECU and to transmit the entire image flow. In fact, the system could just send the alarm, not the entire video stream. In terms of results accuracy, a comparison between the proposed solution and other BSD systems available is not provided in the state of art. The comparison would not be fair because, differently to the presented solution, these systems are usually based not only on motion vector data, but they use other features, as in [15], [16] where authors use edge pixels, lane marks and underneath shadows of vehicles. In [17], a method for detecting the position of a vehicle in the road is presented by using the image data. This method firstly converts the 2D data of a road image into 1D lane information by using the estimation of image entropy and, secondly, the possible vehicle position is determined by differentiating the 1D information between two successive frames. Moreover, the alert is generated by verifying that the detected position lies into a specific area, where different distances from the host car are previously and not automatically measured. In [18], authors present a work for detecting the overtaking vehicle by fusing raw data information from the radar and optical flow from the forward camera.

The rest of the paper is organized as follows: Section II describes the developed algorithm. The Subsections II *A-E* report details of the main steps of the system. Section III shows the performance in terms of complexity and computational cost of the entire system. Three subsections of this paragraph give details about the accuracy of extracted features and about the two main measures performed by the solution. In particular, Subsection III *A* describes a test developed to measure the goodness of the optical flow generated by the system: firstly, the ego-motion is estimated, secondly, it is compared with the real speed of the vehicle. Subsection III *B* shows the assessed accuracy of the vanishing

World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:12, No:7, 2018

point (VP) position. Subsection III *C* reports the assessed accuracy of the overtaking vehicle's distance. Section IV presents conclusions and future developments.

II. ALGORITHM

The proposed method is feature based and the used feature is the Optical Flow (OF), e.g. the collection of Motion Vectors (MVs) indicating the motion of a feature in successive frames.

The OF is computed in hardware ensuring a real time processing (i.e. 30 fps). As shown in Fig. 1, the MVs are the input of the software solution proposed in this paper.

The output is a video showing the overtaking vehicles surrounded by a bounding box colored with three different colors to discriminate different estimated distances. The method starts by filtering the MVs, in particular all the MV are discarded excepting two types: the ones relative to the main road features, as guardrails or lanes, and the ones relative to the incoming vehicles. The first set of vectors is defined as "zoom-out" being features departing from the host car, the second set is defined "zoom-in" representing objects approaching to the host car. The second algorithm's step regards the VP computation based on the set of zoom out vectors. Afterwards, through the usage of VP point and the pinhole camera model, the tilt and the pan camera angles are computed. The aforementioned steps are necessary for computing the calibration camera setting, which is useful to perform an accurate measurement of the distance of the overtaking vehicles. These parameters are temporally computed and tests show a very fast convergence time. The clustering step is devoted to group the zoom-in vectors identifying overtaking vehicles. The next step regards the distance computation of the detected vehicle by using the principles of the pinhole camera model. In the last step, the warning is generated with a risk level proportional to the assessed distance.
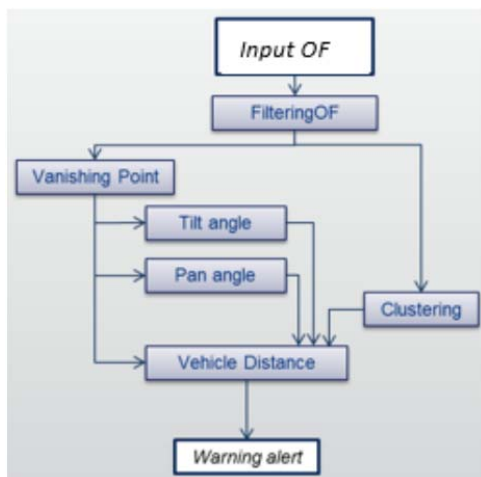

Fig. 1 Algorithm's block scheme

The alert can be of different type as acoustic, visible, haptic or a mixing of them. In this proposal, we realized an output video showing the input scene where the detected overtaking

vehicle is surrounded by a bounding box. According to the distance value changes the color of the surrounding box. Each of the previous mentioned steps is detailed described in the next subsections.

A. OF Filtering

The MVs computed in each frame need to be filtered in order to increase the expected results. In this paper, we propose a filtering based on MVs orientation. Making the assumption that a group of convenient vectors lay into the guardrail position and on the lane markers, a filter based on a set of selected orientations has been applied. The range of orientations has been fixed by allowing the selection of vectors which lay mainly on the road area (lane marker and guardrail). The selected MVs generate the "zoom-out" set. These MVs will be used to compute the VP position in the scene. Another filtering, based on a new set of orientations, is applied to select the vectors representing objects moving toward the host car. These second sets of the selected MV are called "zoom-in" set and it will be used to detect overtaking vehicles. Fig. 2 (a) shows the input OF relative to a given x-frame and Fig. 2 (b) shows the relative extracted set of zoom-out vectors. Fig. 3 (a) shows the input OF relative to another frame and Fig. 3 (b) shows the relative extracted set of zoom-in vectors. The zoom-out will be used to detect the VP position, while the zoom-in will be used to detect the overtaking vehicle.
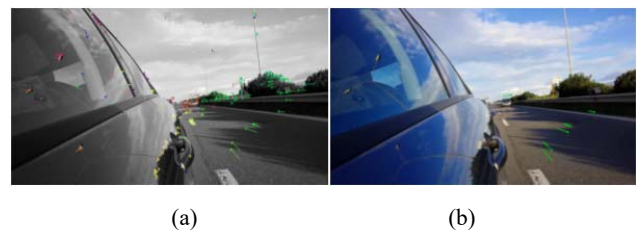

(a)   (b)

Fig. 2 Relative to an x-frame, the image (a) the Input OF and the image (b) the Zoom-Out OF

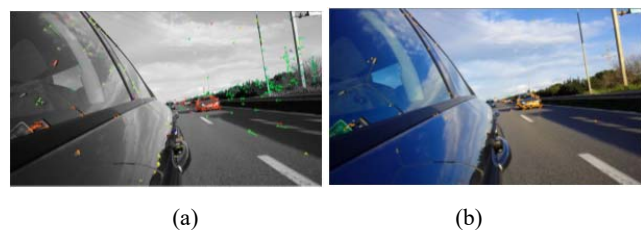(a)   (b)

Fig. 3 Relative to an x-frame, the image (a) the Input OF and the image (b) the Zoom-In OF

B. Vanishing Point (VP)

One of the most useful information for the scene understanding is the position of the VP [19], [20]. In this work, the VP is computed using the OF [21], [22]. In particular, for every MV belonging to the zoom-out set, a line is generated along its direction. Exhaustive intersections among lines are collected. For each frame, the average position of all intersections is performed, and this point is defined as the "current" VP. A temporal averaging is extended

World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:12, No:7, 2018

to subsequent frames, in order to reach a stable VP position. Once a stable position is reached, it will be used to compute two important external camera calibration parameters: the tilt and the pan angles.

The temporal approach allows handling outlier OF, which can be wrongly included into the zoom-out set. The convergence to an accurate value is really fast, tests show that it needs few frames to reach a stable value (more details are reported in Subsection III *B*). The assessed VP position is an important information not only for computing the tilt and pan camera angles, but it is also useful under the scene understanding point of view, allowing to determine the horizon of the scene.

C. Camera Calibration

The computation of the tilt and the pan camera angles lead to the definition of the camera calibration parameters. By changing the camera position in terms of orientation (tilt and pan different to zero), the prospective view of the scene is altered. Considering this real scenario, it is necessary to know the calibration camera setting for assessing a truthful distance measure between the camera and an object into the scene. The basic concepts of the pinhole camera model can be used for measuring the distance between the camera placed on the side mirror of the host car and the overtaking vehicle. In this context, the ideal case must be distinguished, where the

camera has the optical axis parallel to ground and the center of the image corresponds to the VP of the scene, from the real case, where the camera has a tilt angle and a pan angle. Table I shows and explains the meaning of the symbols used for camera calibration.

TABLE I
SYMBOLS FOR CAMERA CALIBRATION

Symbol	Quantity
h	distance of the camera from the ground
f	camera focal length
θ	Tilt camera angle
ψ	Pan camera angle
C	Center of mage
P	Point into the image corresponding to the car from which the distance must be evaluated

Fig. 4 shows the ideal case where a camera is placed on the back of the host camera, with its optical axis aligned with the ground axis (Z). In this case the tilt (θ) and the pan (ψ) camera angles are equal to zero and the center of the image is the same of the VP. Fig. 5 shows the real case where the camera is placed on the side mirror of the host car, with the optical axis not parallel to the Z axis (the ground). In this case, the tilt (θ) and the pan (ψ) camera angles are not equal to zero.
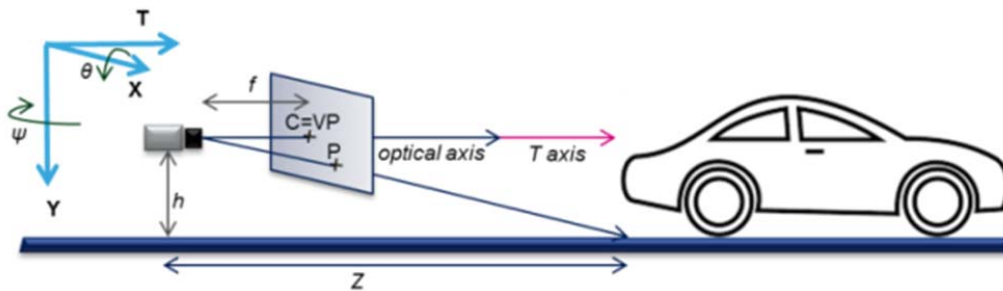

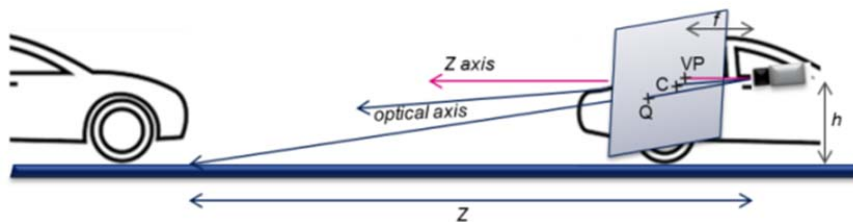Fig. 4 Example showing a camera placed on the back of the host camera

Fig. 5 Example showing the camera placed on the side mirror of the host car

The proposed application is oriented to real applications where the camera calibration step is a common phase to be done. A vertical movement of the optical axis with respect to the ground axis (Z) determines a tilt (θ) angle different to zero, while a horizontal movement of the optical axis with respect to the ground axis (Z) determines a pan (ψ) angle different to zero.

The tilt and pan values are computed by using two formulas derived by mathematical relationships, according of which the

VP lies at a position where the *Z* coordinate of (*X, Y, Z*) approaches infinity, and the relation between pixel and camera coordinates can be derived from the pinhole camera model. Said *VPx* and *VPy*, the VP pixel coordinates and *f* the camera focal length, then the tilt (*ψ*) and the pan (*θ*) angles can be obtained by:

$$\theta = tan^{-1}\left(\frac{VPy}{f}\right) \tag{1}$$

World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:12, No:7, 2018

$$\psi = tan^{-1}\left(\frac{VP_x * cos(\theta)}{f}\right) \qquad (2)$$

The focal length is expressed in pixels, and the conversion from millimeters to pixels is performed as:

$$camF_{pixel} = camF_{millimeters}/factorY \qquad (3)$$

The *factorY* is the ratio between sensor height, expressed in millimeters, and the height of the frame expressed in pixel.

D. Vehicle Detection

Clustering is the task of grouping a set of information in such a way that information in the same group (called a cluster) are more similar to each other than to those in other groups. In this application, it is important to identify moving cars approaching our vehicle. The proposed clustering is based, as mentioned in [23], on only OF analysis. The used system is composed by following basic sub-steps: labelling, clustering, merge clustering. Labelling step inserts labels for each vector. Vectors with the same label are considered to belong to the same cluster. Vectors have the same cluster if they are spatially near and have similar velocity. Clustering step groups vectors with the same label into clusters. At the end, merge clustering step joins clusters with similar orientation, calculated by the media of the vectors inside the considered cluster.

E. Vehicle Distance Assessment

The overtaking vehicles are detected trough the clustering of the zoom-in OF. The proposed method considers the down side of the cluster as the wheel position, on the image plane, of the overtaking vehicle. Through geometrical properties, it is possible to estimate the vehicle's distance from the host car. Taking as reference the setup shown in Fig. 4, where the optical axis is aligned with the ground, the formulas used for detecting the Z value are reported as:

$$f : P = Z : h \qquad \rightarrow \qquad Z = \frac{f \times h}{P}$$

Since the ideal condition showed in Fig. 4 is not usually verified in a real environment, the Z distance value, computed by previous formula, is corrected by the tilt and the pan angles.

F. Warning Alert

The assessed distance is compared with a set of fixed threshold in order to evaluate the level of the risk. In particular, the three distance thresholds (TH) expressed in meters are defined as follows:
1. TH ALERT HIGH RISK   = 5 m;
2. TH ALERT MEDIUM RISK = 10 m;
3. TH ALERT LOW RISK   = 15m

These threshold values can be changed according to the expected sensibility of the system. After the check of the distance value, one of the three levels of the risk is selected:
1. NO_RISK;

2. MEDIUM_RISK;
3. HIGH_RISK.

In the proposed system, the alert is visualized on the output video as a different colored bounding box around the overtaking vehicle: green in case of NO_RISK, yellow in case of MEDIUM_RISK and red in case of HIGH_RISK.

Typically, once the incoming vehicle is detected, the system follows its trajectory during the overtaking by changing the color of the bounding box in the order "green, yellow, and red", till the car disappears from the scene. Fig. 6 shows an example of a vehicle in the overtaking lane far from the host car less than 5 m.

Fig. 6 Example of warning alert; the distance of the overtaking vehicle is lower than 5 m, so the bounding box is signed as red

III. Testing and Performance

The proposed smart side mirror application has been tested with linear lens and using a hardware platform provided by a dedicated chip for the OF detection. Tests were performed also with other commercial cameras, mainly integrated in smartphones. Changing the setup in terms of type of camera, it is necessary to set the relative intrinsic parameters as the focal length and the factorY. It is important also to set the external parameter as the height of the camera with respect to the ground. Knowing these three data, it is possible to compute a single camera parameter which can be computed as follows:

$$CAM\_PARAM = (h \times f) / factorY;$$

Firstly, a video database has been created by using different cameras (all @VGA and @30fps) for collecting more than 5 hours of videos. Running the application on ARM Cortex-A15 r3, single core @ 2.3GHz, the obtained result ensures real time performances (around 100fps). The accuracy of the VP position, the tilt and the pan values and the vehicle distance assessment, have been tested by performing comparison sessions with a ground truth dataset. These evaluations can be surely improved if specialized automotive laboratories are available.

A. Ego Motion Estimation

The goodness of the extracted OFs has been evaluated through a test. First, the ego motion has been computed by using the zoom-out MVs. Secondly, it has been compared to the real vehicle speed directly read from the CAN bus.

The ego motion is estimated for each frame by using the OF laying on the road (the zoom-out OF set). Each single MV represents the distance runs between a frame and the

World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:12, No:7, 2018

successive one. The average of all speeds, one for each MV in the current frame, is used to estimate the velocity of the host car. Frame by frame the velocity can assume different values due to the OF outliers which can wrongly belong to the zoom-out set. The Kalman filter is used to increase the robustness of the assessed speed.

Fig. 7 shows the vehicle speed, one read from CAN bus, and the other one assessed through Kalman temporal filtering. Fig. 8 reports the difference between the two speeds.
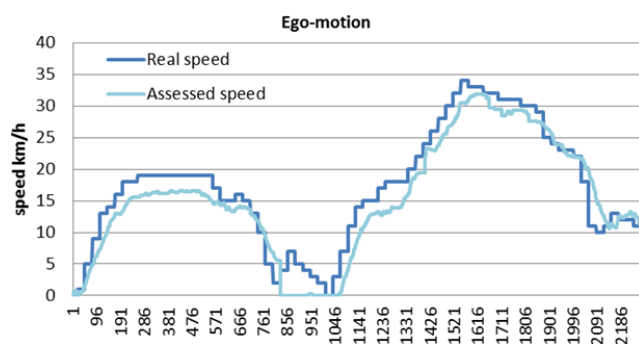

Fig. 7 Comparison between the real speeds of the host car read from CAN bus with respect to the assessed speed computed by using the OF
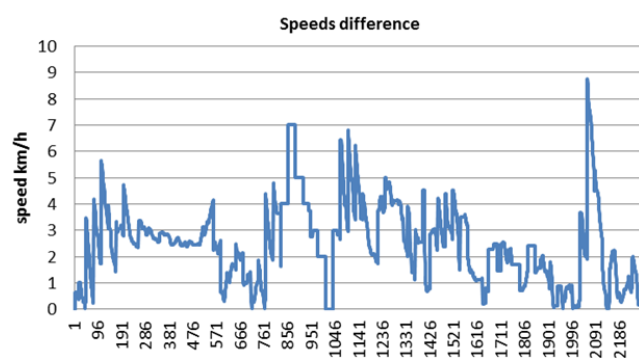

Fig. 8 The absolute value of the difference between the real speed and the assessed one

B. VP Position Accuracy

To evaluate the accuracy of the VP position, a test based on a ground truth comparison has been done. Firstly, the ground truth has been built by annotating manually the VP position for all frames of a set of videos. Then, through the application of the routine described in Subsection II.B, the VP position for each frame of videos used for ground truth production, has been computed and stored in a file. Fig. 9 shows the plot of the VP-x and VP-y coordinates of both the estimated (continuous line) and ground truth values (dotted line) relative to a short test video. It is possible to appreciate the quick convergence of the VP position. In particular, the x and y coordinates reach a stable and a reliable value in less than 30 frames. Repeating the same test for other sequences, comparable results have been obtained.

C. Distance Accuracy

Most efficient methods used to evaluate the distance of an object from the host car are radar based. The proposed solution has not been yet compared to a radar sensor results, but it is one of the next tests to be done. A different approach has been used for testing the accuracy of the assessed distances. Placing the camera in the back car side, a collection of reference points laying on a sort of grid drawn the road have been selected (see Fig. 10). The ground truth has been manually made up by measuring some segments belonging to the grid. A video of few frames of the built scene has been captured by the camera placed on the stationary host car. The tilt and pan angles have been previously computed on other videos acquired with the same camera setup.
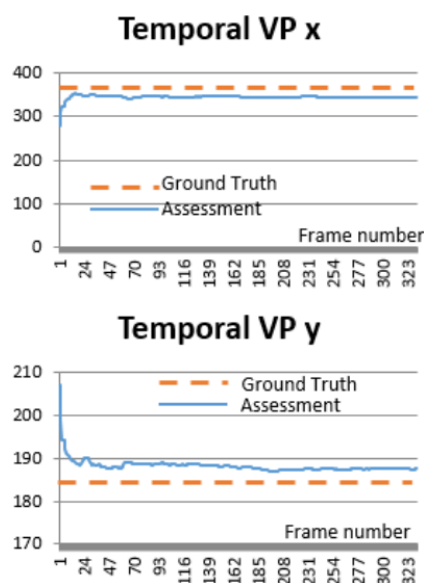

Fig. 9 Representation of the ground truth position of the VP along 323 frames in a video (dotted line), and the VP position assessed through the proposed routine (continuous line)

Fig. 10 Set of reference points used to generate a ground truth

Then, distances between the image plane and the selected points on the image have been computed. In this test, the distance to be evaluated is not relative to the down side of the cluster but it is relative to a set of reference point (see Table II). The tests have been performed on video with a frame size of 640x360. In the table, both the estimated distances and the y coordinates of the points are reported into the scene used for these estimations. The y coordinate is an important value

World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:12, No:7, 2018

because it is directly involved in the formula used for distance computation. The plot in Fig. 11 shows the y position of reference points with respect to the center of the image. Fig. 12 shows the plot of the estimated distances and the ground truth values. It is evident that distances evaluated for points laying too near to the center of the image (as the points E and F in Fig. 10) are less accurate than the distances relative to points far from the image center.
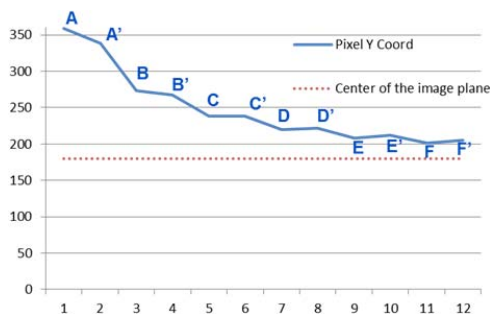
Fig. 11 The y coordinate of the reference points with respect to the center
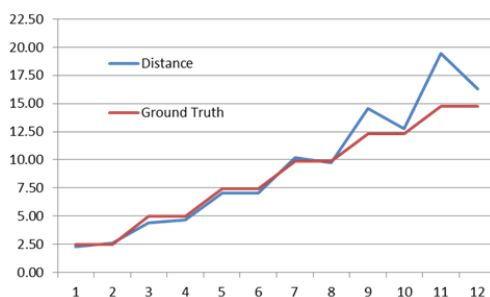

Fig. 12 The assessed distances and the ground truth values

IV. CONCLUSIONS AND FUTURE WORK

TABLE II
DISTANCES: GROUND TRUTH VS. ASSESSMENT

Reference Points	Ground truth (m)	Distance (m)	Pixel y coordinate
A	2.50	2.28	359
A'	2.50	2.58	338
B	5.00	4.39	273
B'	5.00	4.69	267
C	7.40	7.04	238
C'	7.40	7.04	238
D	9.85	10.20	220
D'	9.85	9.72	222
E	12.32	14.57	208
E'	12.32	12.75	212
F	14.77	19.43	201
F'	14.77	16.32	205

Distances relative to a set of reference points.

In this paper, a system for assisting the driver through the monitoring of the back/side of the car relative to an overtaking vehicle is proposed. This application informs the driver in case of the presence of a vehicle into the overtaking lane and gives him information about its distance. Different tests have been performed and results confirm that the proposed solution is suitable for real time systems. Results highlight also the goodness in terms of accuracy of the assessed distances. Many are the developments and possible extensions of the proposed solution. Next tests will regard the evaluation of performances by using fish eye lens together with different camera field of view. Other test will be done to evaluate the accuracy of the estimated distances by comparing the output of the proposed solution with the output of a radar sensor.

REFERENCES

[1] Ismail, M. B., Muqeet, M. A., Malik, M. F. and Khan, A. K. 2014. "Intersection Cross Traffic Warning System for Vehicle Collision Avoidance". In *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* (Vol. 3, Issue 11, November 2014).

[2] Aung, T. and Zaw, M. H. 2014." Video Based Lane Departure Warning System". In International Conference on Advances in Engineering and Technology, Singapore, March 29-30, 2014.

[3] Zhou, Y., Xu, R., Hu, X. F. and Ye, Q. T. 2008. "A Lane Departure Warning System Based on Virtual Lane Boundary". In *Journal of Information Science and Engineering* (24, 293-305, 2008).

[4] Ramesh, S., Ranjan, R., Mukherjee, R. and Chaudhuri, S. 2012. "Vehicle Collision Avoidance System Using Wireless Sensor Networks". In *International Journal of Soft Computing and Engineering*. (2231-2307, Volume-2, Issue-5, November 2012).

[5] Choudhary, A., Moon, Y. and Nawkhare, R. 2015. "Design and Implementation of Wireless Security System in Vehicle". In *International Journal of Innovative Research in Computer and Communication Engineering*. (Vol. 3, Issue 7, July 2015).

[6] Hafner, M. R., Cunningham, D., Caminiti, L. and Del Vecchio, D. 2011. "Automated Vehicle-to-Vehicle Collision Avoidance at Intersections". In *Proc. of ITS World Congress*.

[7] Sotelo, M. A., Barriga, J., Fernández, D., Parra, I., Naranjo, J. E., Marrón, M. and Gavilán, M. 2007. "Vision-based Blind Spot Detection using Optical Flow". In *EUROCAST 2007* (vol. 4739, pp. 1113–1118. Springer, Heidelberg 2007).

[8] Wu, B. F., Kao, C. C., Li, Y. F. and Tsai, M. Y. 2012. "A Real-Time Embedded Blind Spot Safety Assistance System". In *International Journal of Vehicular Technology* (Volume 2012, Article ID 506235, 15 pages). doi: http://doi.acm.org/10.1155/2012/506235.

[9] Scharfenberger, C., Chakraborty, S. and Farber, G.2012. "Robust image processing for an omnidirectional camera-based smart car door". In *ACM Transactions on Embedded Computing Systems* (New York, USA, December, 2012). Volume 11 Issue 4, Article No. 87. doi= http://doi.acm.org/ 10.1145/2362336.2362354.

[10] Gandhi, T. and Trivedi, M. M. 2004. "Motion Based Vehicle Surround Analysis Using an Omni-Directional Camera". In *IEEE Intelligent Vehicles Symposium* (Parma, Italy, June 14-17, 2004).

[11] Paulo, C. and Correia, P. 2003. "Automatic detection and classification of traffic signs". *Proc. 8th Int. Workshop Image Anal. Multimedia Interactive Services* (Santorini, 6-8 June, 2007), doi=http://dx.doi.org/10.1109/WIAMIS.2007.24.

[12] Gavrila, D. M. and Munder, D. M. 2003. "Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle". In *International Journal of Computer* (73(1), 41–59, 2007) doi= http://doi.acm.org/ 10.1007/s11263-006-9038-7.

[13] Sivakumar, R. and Mangalam, H. 2014." RADAR Based Vehicle Collision Avoidance System used in Four Wheeler". In *International Journal of Scientific & Engineering Research* (International Journal of Scientific & Engineering Research).

[14] Händel, P., Yao, Y., Unkuri, N. and Skog, I. 2009. "Far Infrared Camera Platform and Experiments for Moose Early Warning". In *JSAE Trans*, (vol. 40, no. 1, 2009).

[15] Saboune, J., Arezoomand, M., Martel, L. and Laganiere, R. "A Visual Blind-spot Monitoring System for Safe Lane Changes". In *Image Analysis and Processing – ICIAP*, 2011. (2011, pp. 1-10. s).

[16] Din-Chang Tseng, Member, IACSIT, Chang-Tao Hsu, and Wei-Shen Chen. "Blind-Spot Vehicle Detection Using Motion and Static". In *International Journal of Machine Learning and Computing* (Vol. 4, No. 6, December 2014).

[17] Chen, C.-T.and Chen, Y.-S. "Real-time approaching vehicle detection in

World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:12, No:7, 2018

blind-spot area". In *12th International IEEE Conference on Intelligent Transportation Systems.* (pp. 1-6, 2009).

[18] Garcia, F., Cerri, P., Broggi, A., de laEscalera, A. and Armingo, J. M. Data Fusion for Overtaking Vehicle Detection based on Radar and Optical Flow. In *Intelligent Vehicles Symposium.* (Alcalá de Henares, Spain, June 3-7, 2012).

[19] Xiqun, L. 2012. "A New Efficient Vanishing Point Detection From A Single Image". In *IEEE International Conference on Acoustics, Speech and Signal Processing*, (25-30 March, pp. 901-904, 2012).

[20] Kong, H., Audibert, J.-Y.and Ponce, J. 2009. "Vanishing point detection for road detection". In *CVPR,* (pp. 96–103 2009).

[21] DE102014209137 "Method and Apparatus for Calibrating a Camera System in a Motor Vehicle" 2015.

[22] US5777690 A "Device and Method for Detection of Moving Obstacles" 1998.

[23] Spampinato, G., Bruna, A., Curti, S. and D'Alto, V. "Advanced Low Cost Clustering System". In 6th International Conference on Image Processing Theory, Tools and Applications (IPTA) 2016.