Road Detection with Thermal Cameras through 3D Information

G. A. Peláez, D. Bacara, A. de la Escalera, F. García, *Member, IEEE*, C. Olaverri-Monreal, *Member, IEEE*

Abstract— **Ground detection algorithms are important for locating the area where a vehicle can travel during autonomous navigation or for collision avoidance Advanced Driver Assistance Systems (ADAS). Although many algorithms rely on a single camera to detect the ground, they don't consider the possibility of combining transformation approaches with additional information that could be obtained from another camera mounted on the vehicle. In this paper, we propose an algorithm for ground detection that allows the use of disparity maps and the combination of infrared information to enhance road detection. To this end we propose an original approach based on a thermal stereo system to analyze the thermal features of the road that are then complemented by the depth information obtained from the disparity map. We tested the algorithm with a variety of complex scenarios and results showed that the 2D and 3D data enabled accurate detection of the road independent of the weather conditions.**

I. INTRODUCTION

The European Commission's subdivision for Mobility and Transport reports 28,100 road fatalities in the European Union in 2011 alone. This was a 9% decrease from the previous year and a further 8% decrease was expected for 2012. On average, from 2010 to 2013, the European Union has seen a reduction in road fatalities by 17% [1]. In spite of this improvement, new methods and measures for reducing the risk of road users being killed or seriously injured need to be investigated and developed.

One such progressive technology is intelligent vehicles utilizing Advanced Driving Assistance Systems (ADAS) that help the driver to avoid collisions with other road users and objects, such as road lane detectors [2]. They feature obstacle recognition and collision avoidance from other vehicles [3] or pedestrians [4]. In addition, ADAS can help to detect useful information on the road like lanes [5] or road signs [6].

Many of these systems rely on sensors that collect data to identify for instance, the distance to the preceding vehicle, or the information shown on traffic signs. Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication open the possibility of designing Advanced Driver Assistance Systems that use data and information on traffic and road conditions collected by sensors located in other vehicles [5][7],[8],[9] or infrastructure.

G.A. Peláez, F. García and A. de la Escalera are with the Intelligent Systems Laboratory, University Carlos III, Madrid (e-mail: gpelaez@ing.uc3m.es, fegarcia@ing.uc3m.es, escalera@ing.uc3m.es).

D. Bacara is with SIGNUM Bildtechnik GmbH - Rüdesheimer Str. 21 - 80686 München (e-mail: daniel.bacara@signumbt.com).

C. Olaverri-Monreal is with the Austrian Institute of Technology GmbH, Innovation Systems Department, Business Unit Technology Experience, Vienna, Austria (e-mail: cristina.olaverri@ait.ac.at).

Overall, the main idea of the ADAS is to support the driver in the driving task and the decision making process by providing additional information and using a warning system.

Most of these systems alert the driver in case of an imminent unsafe situation [10], [11], [12] and without interfering with the driving task [13]. Additionally, ADAS can help to reduce the number of accidents on the road by predicting unsafe maneuvers by the driver.

This use of this technology is on the rise, and it is expected that by 2025, vehicle-to-vehicle and vehicle-to-infrastructure communication will be available in approximately 30 percent of passenger vehicles on the market, and we will see an increase in the number of autonomous vehicles as well as an improvement in their reliability. By 2030, autonomous-driving-capable vehicles are expected to represent approximately 25 percent of the passenger vehicle population [14]. Therefore, smooth and reliable road detection is crucial for overall road safety in the near future.

Once the road is detected, it can be used by other systems to help the vehicle navigate properly and improve the results of other ADAS on board. As an example, ADAS for pedestrian and obstacle detection can benefit from ground detection algorithms as the road detected can be considered a limited region of interest (ROI) where the ADAS can concentrate its search, thereby reducing execution time and eliminating possible false detections.

This paper proposes a solution for road detection by using the thermal features of the ground. Relying on a pair of thermal cameras (a stereo system) we acquire 3D information from a disparity map and guarantee smoother enhanced road detection compared with approaches based on a single camera.

The next section describes the latest approaches for detecting the ground. In section III and IV, we describe our solution and the testing results in detail so that the advantages of this proposal can be easily noted. The final section discusses our results and the scope of this proposal in certain challenging scenarios.

II. STATE OF THE ART

The earliest publications concerning the problem of road detection date from 1958. A variety of solutions have been researched, with increasing complexity to match improved computational power and processing units, and which have led to a higher quality of results. We would like to highlight some proposals of interest to our current research.

In [15], the authors describe a method composed of three stages: a preliminary classification module, a feature-

 DOI: 10.1109/IVS.2015.7225695

based detection module and an area-based detection module. By combining these three modules, they use image and temporal information with an error correction method. The results are described as robust for the drivable road region.

In a similar group of research, color-based road detection is explained in [16]. As stated in that paper, road detection is one of the key issues of scene understanding for ADAS. The authors propose a solution with a pixel-level confidence map, wherein *v*-disparity map segmentation and an illuminant intrinsic image are utilized. They state that this method performs better in ambiguous environments. The evaluations were executed using a benchmark called KITTI-ROAD.

The authors in [17] describe an approach with a single image without any pre-learned information. They propose a graph-cut segmentation framework, whereby the training data is produced from an estimated ground region of the current frame. They perform this method repeatedly until an accurate region is obtained. The results are compared to the SUN Database.

In [18] a method that with road models is shown. Based on road classification, it decreases the chance that an analyzed frame contains certain types of erroneous features. In their work the authors argue that the classifier learns a probability map for each type of geometry during the training stage. Their experiments demonstrate that the proposed solution outperforms some of the latest algorithms in the context of a frame-by-frame analysis.

A different proposal using Support Vector Machines (SVM) is proposed by the authors in [19]. Their main goal is to correct problems of feature extraction and classification by using SVM for supervised learning. According to the authors the solution is capable of updating the data for online training in order to reduce the possibility of incorrect classifications. It is stated to be a novel framework for self-supervised online learning.

Alhwarin et al explain in [20] an improvement on depth images using structured light from devices with Stereo IR projectors. They show the results in different formats including the construction of clouds of points. It is oriented to indoor applications and it is not adequate for locating the area where a vehicle can travel.

In line with our stereo system, the authors in [21] compare infrared and color stereo systems oriented to pedestrian detection. By combining *u* and *v* disparity with ground plane estimation they obtain a bounding box for the detected pedestrians. Results showed a high level of accuracy when detecting the pedestrians with both systems.

A nonparametric method is proposed in [22] where the intrinsic parameters of the road are used. Different scenarios were tested to detect the road with sequences from different publishers. Additionally, they removed some obstacles by analyzing only the *v* disparity map.

Finally, [23] explains how to process dense stereo data by using elevation maps. A digital elevation map (DEM) is generated according to a mathematical estimation of a road. By comparing different points of this map to the DEM, it can be distinguished if the analyzed region is an obstacle or part of the road, similar to the *v* disparity map analysis from [22]

The approach in our proposal sets itself apart from others in that it takes into account the thermal information from one camera and complements it with results obtained from analyzing the disparity map from a stereo thermal camera system. This results in a robust approximation of the road i.e. it will detect an object that has a regular thermal fingerprint and no texture that could return any information to the disparity map. This is a real time process.

III. System Description

The proposed solution receives images from a stereo thermal camera system embedded in a vehicle that is an experimental platform for multiple Advanced Driving Assistance Systems. From this set of thermal sensors, two images are obtained. These images are subject to different image processing algorithms such as erosions, thresholds, and stereo matching.

In this section we describe the process in detail. The first algorithm to be executed analyzes the left image obtained from the stereo thermal set and compares a reference value with the lower part of the image. The goal is to find all areas that have a similar chromatic feature to the road as described below:

A. Chromatic similitude evaluation

Assuming that the left camera will always be located in the same position, we presumed that the pixels in the inferior part of the image belong to the ground i.e. road detection. From this inferior region, the average thermal value is measured and it is used as a reference, with a specified error tolerance, to analyze the rest of the image by comparing it with the thermal information available from one single frame of the left camera.

A matrix *Mg*(1) containing thermal values is built with those pixels with thermal information similar to the reference temperature, as seen in equation (1). The result is a mask with white pixels where the thermal value (*Gv*) is between the range of what is considered the reference value (*Gr*) and an empirically calculated tolerance (σ).

The value of the pixel at the coordinates *m,n* at equation (1) will be included in a matrix *Mg*(2) with the same dimensions as the input image if the measured value of the pixel is between the limits defined by the reference Gr value and the tolerance σ. For the ease of result analysis, the pixels that pass this criterion are painted in white to create a mask.

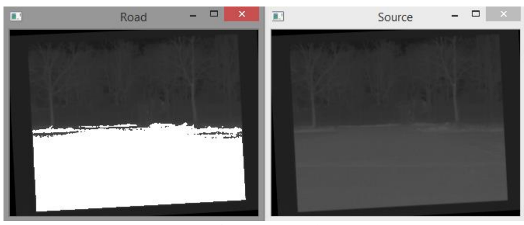
$$\mathrm{Mg}=\begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{m1} & \cdots & P_{mn} \end{bmatrix} \tag{1}$$

$$P_{mn} = \begin{cases} 0, & |Gv(\mathrm{m},\mathrm{n}) - Gr| < \sigma \\ 255, & |Gv(\mathrm{m},\mathrm{n}) - Gr| \geq \sigma \end{cases} \tag{2}$$

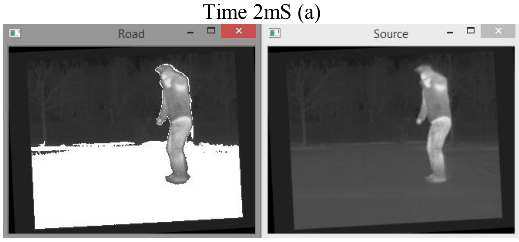
The described algorithm is intended to be exported from a CPU computation to a GPU parallel computation. Because it does not depend on the previous analysis of the pixels, this algorithm is executed faster than other solutions tested. For example, other algorithms sweep the image and compare the difference of one pixel with the previous one, and if the difference is too great, the upcoming/following pixels would not be considered as part of the ground. This type of approach had the drawback of dependency on the analysis of the previous pixels, slowing parallel execution. Chromatic similitude algorithm does suffer from this drawback. The left column of Fig. 1. shows the rectified image of the left camera and next to it the result of the sampling with the time of execution. All images have the same dimensions (320x240).

B. *Filtering the results with disparity maps*

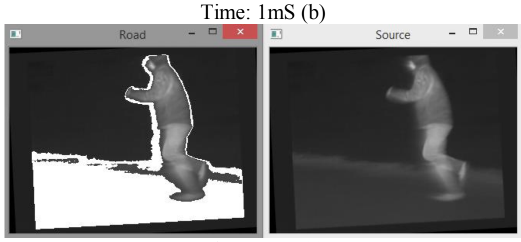
As the approach described above relies on a thermal camera stereo system, it is possible to build a disparity map of the scene, and obtain 3D information which can support the 2D (chromatic features) algorithm.



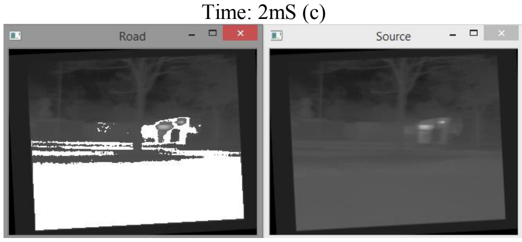
Time 2mS (a)



Time: 1mS (b)



Time: 2mS (c)



Time: 2mS (d)

Time: 2mS (e)

Figure 1(a) to (e): Execution time and results of the chromatic algorithm. The images on the left image show the results in white. The images on the right are the source images that have not yet been processed.

It is taken as a prerequisite for this approach that the surface labeled as road is regular without significant bumps or holes, as these would be considered obstacles.

The fact that the road to be analyzed will not always have a homogeneous temperature must be considered. This is due to the presence of shadows of objects such as trees or buildings, as well as the presence of puddles or different types of asphalt.

The chromatic approach, based on thermal information, is similarly unable to identify such areas as belonging to the road. Therefore, a simple increment of the range of values to be considered as valid by the chromatic algorithm may help to identify some or all of these pixels as belonging to the road, but also may lead to falsely attributing obstacles to the road map as seen in Fig. 2.

In order to solve this problem, the disparity map obtained from the stereo system is used. The disparity map allows to calculate the distance of an object to the test vehicle by taking into account the disparity of the location of the pixels corresponding to the same object in the two images (3).

$$distance = \frac{f \cdot B}{u_L - u_R} = \frac{f \cdot B}{d} \tag{3}$$

where d is the disparity in pixels, f is the focal length in mm per pixel and B is the baseline (distance between the stereo cameras).

For this calculation both images should be coplanar, their optical axes must be parallel, the intrinsic parameters should be equal and the correspondence between points in both images has to be found, in order to determine the disparity. This means that those objects that were detected by the chromatic algorithm such as lamp posts, fences, trees etc… are shown in the disparity map with their respective depth. However, in order to calculate the disparity, the obstacles detected rely on visual features, such as border, corners... due to the lack of texture given by the absence of thermal gradient, the part of the image that belongs to the road will not be displayed in the disparity. In these situations, the Stereo Block Matching (SBM) algorithm used on this approach cannot obtain matching textures, therefore no disparity information is provided. The parameters for the SBM algorithm were established from the analysis of the results obtained both for the road detection and other algorithms that required the 3D information.

In the next process, a third image is created, *Md* (4) composed by the pixels *Q* with coordinates *m,n* with the value defined in (5). This image will show those regions that have no texture i.e. without information enough to create the disparity.

$$\text{Md}=\begin{bmatrix} Q_{11} & \cdots & Q_{1n} \\ \vdots & \ddots & \vdots \\ Q_{m1} & \cdots & Q_{mn} \end{bmatrix} \tag{4}$$

$$i = 1..m; j = 1..n;\ Q_{ij} = \begin{cases} 0, & d(\mathrm{i,j}) > 0 \\ 255, & d(\mathrm{i,j}) = 0 \end{cases} \tag{5}$$

where $d(m,n)$ corresponds to the disparity value, of the disparity map in the coordinates *m,n*.

It can be assumed that the only object located in the lower part of the image with a regular surface without a relevant texture is the ground i.e. the road. Therefore, the area that has no information in the disparity map belongs to the road. Sidewalks, grass, parked vehicles and pedestrians will have a relevant texture that SBM can use to return a valid value. The parameters for the SBM algorithm (specifically the texture threshold) were determined from the analysis of a number of frames from sequences recorded during winter and summer. After the disparity map is obtained, an opening is performed on it so that the dark regions can be enhanced and the noise from random white pixels can be eliminated.

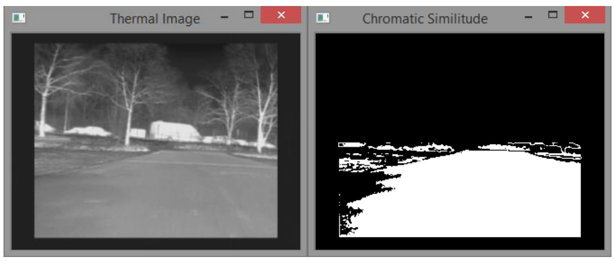


Figure 2: Example of a non-regular surface with noise included

With the obtained result of the disparity map (Fig. 3), the chromatic algorithm can be refined. This is done by considering only those pixels in the disparity map which make up the dark area without information. As a result, the image *Mr* (5) is built with pixels whose value *V3*(6) depends on the two previous images (disparity and chromatic). Figure 5 shows the graphical representation of the proposed solution. The first part is a masking process that filters those pixels that are not close to a reference value. The second part filters these resulting pixels by applying another mask that takes into account only those that belong to an area with no texture.

$$\mathrm{Mr}=\begin{bmatrix} I_{11} & \cdots & I_{1n} \\ \vdots & \ddots & \vdots \\ I_{m1} & \cdots & I_{mn} \end{bmatrix} \tag{5}$$

$$\mathrm{I_{mn}} = \mathrm{P_{mn}} \cap \mathrm{Q_{mn}} \tag{6}$$

IV. Algorithm Testing Results

The result from this masking process is an image that takes into account only those pixels that have a chromatic similitude to the segment of the ground in front of the vehicle and that belong to a smooth regular surface (the road). The pixels detected as being part of the road were marked in white. During the chromatic similitude approach many of the pixels belonging to the road were distinguished, but also included were some that should not have been. During the second approach, these wrongly detected pixels were mostly erased after the depth information from the disparity map was taken into account. Although it filtered the image, it also removed some pixels that should have remained. However, overall, a large region that belongs to the road was detected successfully. Finally, with the erosion and dilation executed on the image, the detected object was more regular i.e. less black pixel areas surrounded by white. The whole algorithm is visualized in Fig. 4 and Fig. 5.

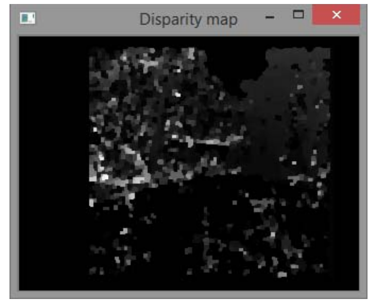


Figure 3: Example of a disparity map obtained from the stereo thermal cameras using the SBM algorithm

To test the reliability of the algorithm, a set of images was selected during different weather conditions. This set of images is composed by the most relevant from a sequence containing more than 3800 for summer and winter. Due to the high frame rate of the capturing process, consecutive images do not vary significantly one from another thus only the most different were chosen from the sequence. In this case, 25. Then, these chosen images were manually marked for regions that belong to the road. The algorithm analyzed the original images and compared the position of each pixel considered as belonging to the road with the marked images. After 25 different images of the road in winter, the average percent of the road detected was 96.28%. Another 25 images of the road during summer were used, and here the average was 83.71%. In Fig. 7, the algorithm marks the detected road in white.

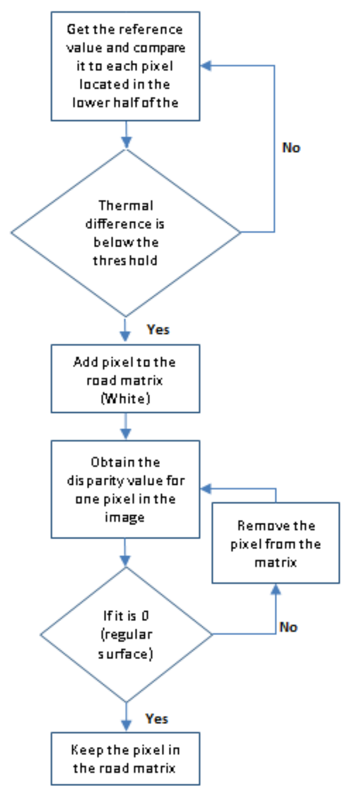


Figure 4: Block diagram of the proposed solution

Figure 5: Overall example of the proposed solution.

The pixels that should be part of the road but were rejected by the algorithm because they didn't fit the different criteria are marked in light grey. Table I depicts some of the test results for 3 images in winter and 3 images in summer. This selection includes the worst detection rate that occurred within the 50 images tested (Fig. 6f). All results yielded a rate higher than 50% positive road detection (Table I).

	Name	***Ground Truth Pixels***	***Road Detected Pixels***	***Percent of Positive Pixels Detected***	***Time (mS)***
Winter	Fig. 6.a	21284	18169	85.36	3
	Fig. 6.b	21570	20696	95.95	5
	Fig. 6.c	24117	23951	99.31	4
Summer	Fig. 6.d	12623	10554	83.61	5
	Fig. 6.e	20683	19066	92.18	3
	Fig. 6.f	20019	11401	56.95	5

TABLE I. RESULTS OF THE PROPOSED ALGORITHM WITH 3 SUMMER AND 3 WINTER SITUATIONS

V. CONCLUSION AND FUTURE WORK

Further research will emphasize improvement of the parameters for both chromatic and SBM algorithms in order to strengthen their results in scenarios that currently present a challenge for the approach presented in this work. Also of interest is the determination of the equation of the plane that fits the points that belong to the road. This could be done in the event that there are other sensors that could deliver more information about the roadway i.e. to obtain a depth map with valid points. To prove that it is possible to determine the plane equation of the road, an artificial image (computer generated) was used as an example in our approach (Fig. 6). Notice that the floor has a texture and thus the determination of the plane is possible thanks to proper results obtained from the depth map.

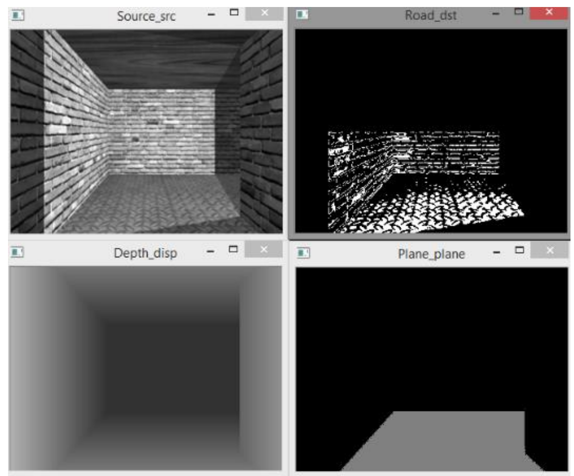


Figure 6: Artificial image to test the plane fit equation approach

ACKNOWLEDGMENT

This work was possible thanks to the combined collaboration of Signum GmbH and the Institute of Ergonomics at the department of Mechanical Engineering, Technische Universität München. Additionally, this work was supported by the Spanish Government through the CICYT projects (TRA2013-48314-C3-1-R) and (TRA2011-29454-C03-02).

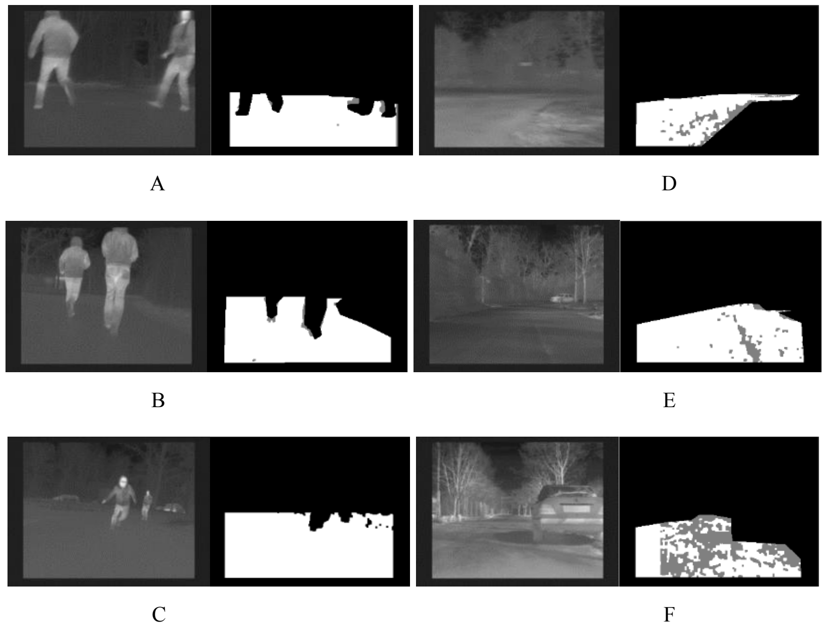


Figure 7: Visualization of the results. The estimated pixels as being the road are in white, the (real road, ground truth, reference, mask) are in gray.

REFERENCES

[1] European Comission, division of Mobility and Transport. “Road Safety evolution in the EU” Statistical Figures, March 2014. http:// http://ec.europa.eu/transport/road_safety/specialist/statistics/index_en.htm

[2] J. M. Collado, C. Hilario, A. de la Escalera, J. M. Armingol Multi-lane visual perception for lane departure warning system. Third International Conference on Computer Vision Theory and Applications, pp: 360-367 (2008)

[3] B. Musleh, A. de la Escalera, J. M. Armingol. U-V Disparity Analisys in Urban Environments. 13th International Workshop on Computer Aided Systems Theory, pp: 169-170 (2011).

[4] Soquet, N., Perrollaz, R., Labayrade, R., Auber, D.: Free Space Estimation for Autonomous Navigation. 5th International Conference on Computer Vision System, Bielefeld, pp. 1–6 (2007).

[5] U. Franke, D. Gavrila, A. Gern, S. Gorzig, R. Janssen, F. Paetzol, C. Wohler. From door to door, principles and applications of computer vision for driver assistant systems,” in Intelligent Vehicles Technologies, L. Vlacic, F. Harashima, and M. Parent, Eds. Oxford, U.K.: Butterworth Heinemann, pp. 131–188 (2001).

[6] G. A. P. Coronado, M. R. Muñoz, J. M. Armingol, A. de la Escalera, J. J. Muñoz, W. van Bijsterveld, and J. A. Bolaño, “Detection and classification of road signs for automatic inventory systems using computer vision,” Integr. Comput. Aided. Eng., vol. 19, no. 3, pp. 285–298, Jan. 2012.

[7] C. Olaverri-Monreal, P. Gomes, R. Fernandes, F. Vieira, and M. Ferreira, “The See-Through System: A VANET-enabled assistant for overtaking maneuvers,” in Proceedings of the IEEE Intelligent Vehicles Symposium (IV), 2010, pp. 123–128.

[8] P. Gomes, C. Olaverri-Monreal, and M. Ferreira, “Making vehicles transparent through v2v video streaming,” IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 2, pp. 930–938, 2012.

[9] C. Olaverri-Monreal, A. Hasan, J. Bulut, M. Körber and K. Bengler, "Impact of In-Vehicle Displays Location Preferences on Drivers’ Performance and Gaze.” In IEEE Transactions on Intelligent Transportations Systems (IEEE T-ITS), vol. 15, issue 4, pp. 1770 – 1780, 2014..

[10] F. García, J. García, A. Ponz, A. de la Escalera, and J. M. Armingol, “Context aided pedestrian detection for danger estimation based on laser scanner and computer vision,” Expert Syst. Appl., vol. 41, no. 15, pp. 6646–6661, 2014.

[11] D. Martín, F. García, B. Musleh, D. Olmeda, G. Peláez, P. Marín, A. Ponz, C. Rodríguez, A. Al-Kaff, A. de la Escalera, and J. M. Armingol, “IVVI 2.0: An intelligent vehicle based on computational perception,” Expert Syst. Appl., vol. 41, no. 17, pp. 7927–7944, Jul. 2014.

[12] G. A. Pelaez C, F. Garcia, A. de la Escalera, and J. M. Armingol, “Driver Monitoring Based on Low-Cost 3-D Sensors,” Intell. Transp. Syst. IEEE Trans., vol. 15, no. 4, pp. 1855–1860, 2014.

[13] F. García, F. Jiménez, J. J. Anaya, J. M. Armingol, J. E. Naranjo, and A. de la Escalera, “Distributed pedestrian detection alerts based on data fusion with accurate localization.,” Sensors (Basel)., vol. 13, no. 9, pp. 11687–708, 2013.

[14] http://www.gartner.com/newsroom/id/2884317. Accessed: 18th January 2015,

[15] Chunzhao Guo; Mita, S., "Drivable road region detection based on homography estimation with road appearance and driving state models," Autonomous Robots and Agents, 2009. ICARA 2009. 4th International Conference on , vol., no., pp.204,209, 10-12 Feb. 2009

[16] Wang, B.; Fremont, V.; Rodriguez, S.A., "Color-based road detection and its evaluation on the KITTI road benchmark," Intelligent Vehicles Symposium Proceedings, 2014 IEEE , vol., no., pp.31,36, 8-11 June 2014

[17] Zhen He; Tao Wu; Zhipeng Xiao; Hangen He, "Robust road detection from a single image using road shape prior," Image Processing (ICIP), 2013 20th IEEE International Conference on , vol., no., pp.2757,2761, 15-18 Sept. 2013.

[18] Alvarez, J.M.A.; Gevers, T.; Lopez, A.M., "Vision-based road detection using road models," Image Processing (ICIP), 2009 16th IEEE International Conference on , vol., no., pp.2073,2076, 7-10 Nov. 2009

[19] Shengyan Zhou; Jianwei Gong; Guangming Xiong; Huiyan Chen; Iagnemma, K., "Road detection using support vector machine based on online learning and evaluation," Intelligent Vehicles Symposium (IV), 2010 IEEE , vol., no., pp.256,261, 21-24 June 2010.

[20] F. Alhwarin;, A. Ferrein; I. Scholl; “IR Stereo Kinect: Improving Depth Images by Combining Structured Light with IR Stereo” Proc. of the 13th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2014), vol. 8862, no. 2014, pp. 409-421

[21] Krotosky, S.J.; Trivedi, M.M., "A Comparison of Color and Infrared Stereo Approaches to Pedestrian Detection," Intelligent Vehicles Symposium, 2007 IEEE , vol., no., pp.81,86, 13-15 June 2007

[22] Wu, M.; Lam, S.; Srikanthan, T., "Nonparametric Technique Based High-Speed Road Surface Detection," Intelligent Transportation Systems, IEEE Transactions on , vol.16, no.2, pp.874,884, April 2015

[23] Oniga, F.; Nedevschi, S., "Processing Dense Stereo Data Using Elevation Maps: Road Surface, Traffic Isle, and Obstacle Detection," *Vehicular Technology, IEEE Transactions on* , vol.59, no.3, pp.1172,1182, March 2010