2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Obstacle Detection and Localization Method Based on 3D Model: Distance Validation with Ladar

Cindy Cappelle, Maan El Badaoui El Najjar, François Charpillet and Denis Pomorski

Abstract— **This paper presents a method for the detection and the geo-localization of obstacle in outdoor environment. The proposed approach exploits a geographical 3D model managed by a 3D Geographical Information System (3D-GIS) and a video camera. An image processing module is developed to match synchronized real image and virtual image. The real image is provided by an on-board camera and gives the real view of the scene seen by the vehicle whereas the virtual image is provided by the 3D-GIS. A centimetric LRK GPS provides an estimation of the geo-position of the vehicle, which is used to localized the vehicle in the geographical 3D model. Obstacle(s) are then detected and tracked by comparison between the real image that contain the obstacle(s) and the virtual image where obstacle(s) are absent. The developed method permits also to compute the distance between the camera and obstacle(s), as well as the geo-position of these detected obstacles. In order to validate the computed distance, a 2D ladar (laser range scanner) is used. A necessary condition to complete the comparison is to calibrate the camera and the ladar together. Experimental results with real data are presented in the final section.**

I. Introduction

OBSTACLE detection is a key component of an autonomous vehicle and develops more and more for driver assistance [1], [9]. A wide range of sensors and various methods for detecting obstacles have been proposed. The used sensors are mainly: radar [16], laser [7], [12], monocular vision [3], [8] and stereovision [2], [5], [10], [17] or multisensor fusion [6].

In this work, an obstacle detection and geo-localization method using a geographical 3D model and a camera is presented. The vehicle is first localized in the 3D model through the position given by the centimetric LRK GPS. Using this GPS position, the snapshot of the virtual camera is provided by the 3D-GIS. The virtual image and the real one, captured by an on-board camera, are then treated to detect the static and dynamic obstacles in the scene. The detected obstacles are finally geo-localized using the depth image also provided by the 3D-GIS in addition to the virtual image. The obtained results are validated with a ladar.

Manuscript received September 14, 2007.

C. Cappelle, M. El Badaoui El Najjar, and D. Pomorski are with the Laboratoire d'Automatique, Génie Informatique et Signal, Polytech'Lille, 59655 Villeneuve d'Ascq Cedex, France (phone: +33 (0)320434659 ; fax: +33 (0)320337189 ; e-mail: Cindy.Cappelle@polytech-lille.fr, Maan.E-elnajjar@univ-lille1.fr denis.pomorski@univ-lille1.fr)

F. Charpillet is with the Laboratoire Lorrain de Recherche en Informatique et ses Applications, Campus Scientifique, BP 239, 54506 Vandoeuvre-lès-Nancy, France (e-mail: francois.charpillet@loria.fr).

This paper is organized as follows. In Section II, the used information sources are presented: the geographical 3D model managed by 3D-GIS and the on-board sensors (camera, GPS, laser scanner). Section III describes our method, which uses geographical 3D model, camera and GPS, to detect and geo-localize obstacles. Section IV presents the experimental results obtained with our method, and compares these results with ladar measurements. Finally, section V concludes the paper.

II. The different information sources

The proposed method uses an accurate geographical 3D model managed by a 3D Geographical Information System (3D-GIS) and various embedded sensors: a camera, a GPS receiver and a laser scanner. This section describes these different information sources and how to combine them.

A. Geographical 3D model and 3D-GIS

To manage in real time data from a geographical 3D model (also called 3D geographical database, 3D city model or 3D virtual model), a 3D-GIS is required. A 3D-GIS is a computer system capable of integrating, storing, editing, analysing, sharing and displaying geographically-referenced information. In a more generic sense, GIS is a tool that allows both to create interactive queries (user searches) and to analyse the spatial information, and edit data [14].

Such 3D models provide a huge potential to vehicle localization and navigation applications [11], [13]. For example, matching 3D data set with video stream allows to determine the position and orientation of a human or a vehicle in urban environments. All major Japanese cities have been covered since 2002 and are updated every six months. The French project Bati3D is a very important project which is aiming at developing a 3D geographical map for the major cities in France. Laser profiler data, 2D digital map and aerial image are used to generate automatically the 3D geographical model [15], [18].

The used 3D-GIS

Fig.1: 3D images of the Stanislas Square in Nancy

The used geographical 3D model (Fig.1) is provided by the Tecnomade[1] company and has a metric accuracy. The Tecnomade 3D engine, which can manipulate the 3D geographical model format of Tecnomade, was used to develop our 3D-GIS, adapted to application in robotics and intelligent vehicle. For the development of the proposed obstacle detection and geo-localization approach, several functionalities have been developed and integrated to the 3D-GIS. The two functions used in this work are illustrated on Figure 2. The first one lets extract a virtual image in Bitmap format, by giving to the 3D-GIS the 6 degrees of liberty of the virtual camera in the local coordinates system of the geographical 3D model (Fig.2.a). The second one provides in a binary text file the depth of each virtual image pixel (Fig.2.b). In this file, distances in meter between the virtual camera and the 3D point represented by each pixel are stocked in 32 bits long format. The depth information is extracted from the Z-buffer of the video card. A more complete description (with all the developed functionalities) can be found in [4].

a

8.8784628e+000	8.8781366e+000
8.8778124e+000	8.8779469e+000
8.8776264e+000	8.8773079e+000
8.8769913e+000	8.8771334e+000
8.8768196e+000	8.8765087e+000
8.8766565e+000	8.8763485e+000
8.8760433e+000	8.8757391e+000
8.8758945e+000	8.8755941e+000

b

Fig.2: The used 3D-GIS output functionalities

B. *On-board sensors*

1) Camera

The AVT MARLIN F-046C camera, connected on the FIREWIRE IEEE 1394 SVGA bus, provides 640x480 pixels images at the acquisition frequency of 30 Hz. As all vision system that extracts accurate geometrical information of a scene, the camera was calibrated [19], [20]. The intrinsec parameters (focal, principal point, scale factor) and the distorsion parameters was thus determined.

The role of camera in our method is to provide a real view of the vehicle field of vision, i.e. the scene with the possible obstacles. These obstacles can then be detected by comparing this real view (including obstacles) with the virtual image extracted from the 3D model i.e. the scene without obstacle.

[1] www.tecnomade.fr

2) GPS

The Sagitta02 GPS receiver supplied by Thales Navigation provides vehicle geo-position whose real-time precision ranges from the meter to the centimeter level, depending on how it is operated: GPS, WAAS/EGNOS, DGPS or LRK mode. Its acquisition frequency was fixed at 10 Hz.

The GPS function is to provide a pose estimation of the vehicle to the 3D-GIS. A function was developed to convert the position from the GPS coordinates system (WGS84) to the 3D model local coordinates system. This position permits to place the virtual camera in the 3D model, the 3D-GIS extracts then the virtual view seen from this location.

3) Laser scanner

The SICK LMS 291 laser scanner can be used for standard applications involving measurement of objects and position determination, monitoring areas, vehicle guidance and collision control. The laser scanner provide accurate distance measurement throughout the 180° scanning field with 1° angular resolution. Ladar measurements are based on time-of-light measurement. The laser scanner calculates the distance to the object using the time-of-light, i.e. the length of time between sending and receiving the beam of light.

The laser scanner is used to validate the obstacle detection and geo-localization results we obtained with vision and 3D model.

Those three sensors were fixed on the front of the vehicle as shown on Figure 3.

Fig.3: Sensors embedded on the vehicle

III. Obstacles Detection and Geo-Localization

Figure 4 describes the obstacle(s) detection and geo-localization method. The application principle is to compare two images: the real image where possible obstacles appear and the virtual image where obstacles aren't. The real image is provided by an on-board camera whereas the 3D-GIS extracts the virtual image. In order to return the snapshot, the 3D-GIS needs the vehicle pose as input. The 3 degrees of liberty (X, Y and yaw) of the virtual camera (the roll, pitch and elevation were fixed for the experimental area) are obtained from the conversion of GPS measurements to the local 3D model frame. LRK GPS provides effectively longitude and latitude with centimetric accuracy of the vehicle in the *WGS84* coordinates system and the heading θ of the vehicle. The real and virtual images can then be treated to extract the obstacles in the scene. With the integration of the depth information of 3D-GIS in the image processing algorithm, the distance between the detected obstacles and the vehicle is also provided. Finally, the detected obstacles are geo-localized in the *WGS84* coordinate system.

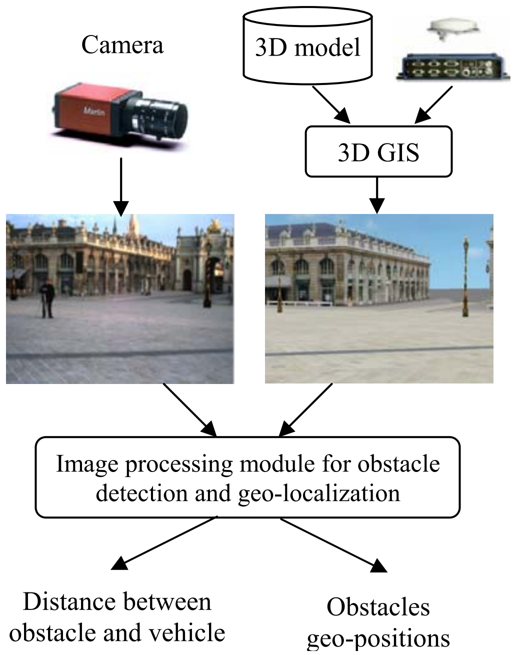

Fig.4: Obstacles detection and geo-localization method

A. Image processing for obstacle(s) detection

The aim of this part is to describe how obstacle(s) are detected at each instant *k* using the two considered images (Fig.5):

- the real image: image captured by the on-board camera,
- the virtual image: image provided by the 3D-GIS.

To have an accurate obstacle(s) detection and geo-localization, the virtual and real camera have to be synchronised spatially and timely. For spatial synchronisation, the geo-position and the orientation of the virtual camera in the 3D model are computed using the centimetric LRK GPS (Sagitta 02 Thales product). The GPS antenna is mounted near the real camera in order to have the accurate position of the real camera. Furthermore, for time synchronisation, the real and virtual images are choosen so that the acquisition dates of the real camera and the GPS correspond.

Fig.5: real (left) and virtual (right) images at the same instant k

As the obstacle detection method is developed to be implemented in real time application, a simple image processing algorithm is required. After a pre-treatment stage, an image processing method based on the computing of the difference between the binarised virtual image and the binarised real image is developed. Note that to avoid the detection of elements in the model which aren't in the real scene, an order in the substraction has to be respected. Moreover, as the upper part of the image doesn't contain useful information about the obstacle(s), the only lower part is conserved. Finally, a post-treatment is realised, in order to eliminate the noise. The results of these different steps are illustrated on an example in Figure 6.

As we supposed the obstacles are constrained to the ground plane, the lowest pixel of each detected obstacle in the difference image is the intersection between the floor and these obstacles.

Fig.6: Obstacle detection result

B. Computation of vehicle/obstacle(s) distance

In the previous part, the obstacle(s) were detected and the obstacle(s) pixels were isolated. The distance to this/these obstacle(s) has now to be computed. In order to achieve this aim, the depth information from the Z-buffer is used. The depth information of each pixel of the virtual image is provided by the 3D-GIS in an binary file. As the real and the virtual images have the same projection plane, the real obstacle(s) can be projected in the depth virtual image in order to extract the depth of pixels representing obstacle(s) (Fig.7).

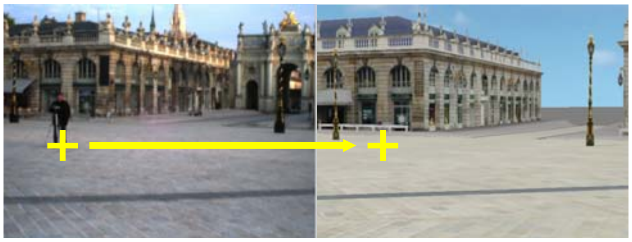

Fig.7: Real and virtual images matching

The depth (denoted *l)* of the pixels which correspond to the detected obstacle(s) is extracted from the binary depth file provided by the 3D-GIS. It gives the distance between the camera and the obstacle(s). The error announced by the geographical 3D model constructor is under 10 cm.

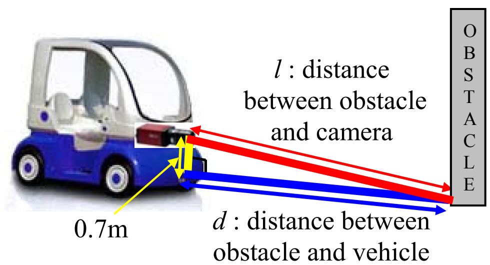

Fig.8: Distance computing between obstacle and vehicle

The camera was installed on the front of the vehicle, as shown in Figures 3 and 8. So the distance (denoted *d*) between the front of the vehicle and the obstacle is defined as follows:

$$d = \sqrt{\left|l^2 - 0.7^2\right|}$$

C. Detected obstacle(s) geo-localization

Once the distance between vehicle and obstacle is calculated, the obstacle can be geo-localized. Since the height of the obstacle isn't needed, the scene can be projected on the *XY* plane, which is supposed to be the floor plane. The obstacle belongs then to the line which passes through the camera and which equation is given below:

$$y = cotan(\beta).(x - X_C) + Y_C$$

with:

- X_C and Y_C, the camera coordinates in the *Lambert93* system (given by LRK GPS),
- β, the orientation of the obstacle in the *Lambert93* system, i.e. the angle between the obstacle and the North axis.

As the distance between the camera and the obstacle is also known in the geographical *Lambert93* system, the obstacle is then geo-localized.

The procedure to calculate the β parameter is now explained. Let's denote $head$ the vehicle orientation i.e. the angle between the camera axis and the Y axis (North) of the *Lambert93* system and α the angle between the camera axis and the obstacle. The orientation β is then:

$$\beta = head - \alpha$$

The angle α formed by the orthogonal axis of the image plane at the image center (segment *D* in Figure 9) and the segment d_{ob} linking the camera to the obstacle has to be computed. In Figure 9, one can see that in the same depth plane, the angle α is calculated using the following trigonometric equation:

$$\tan\alpha = \frac{A}{width/2}\tan(O_h/2)$$

where:

- O_h represents the horizontal Field Of View (FOV) fixed in the 3D-GIS,
- $width$ is the horizontal pixels number of the image fixed by the used image resolution,
- A represents the pixels number between the pixel of the center of the image and the projection P_{ob} of the obstacle on the plane P whose depth is the depth of the central pixel of the image.

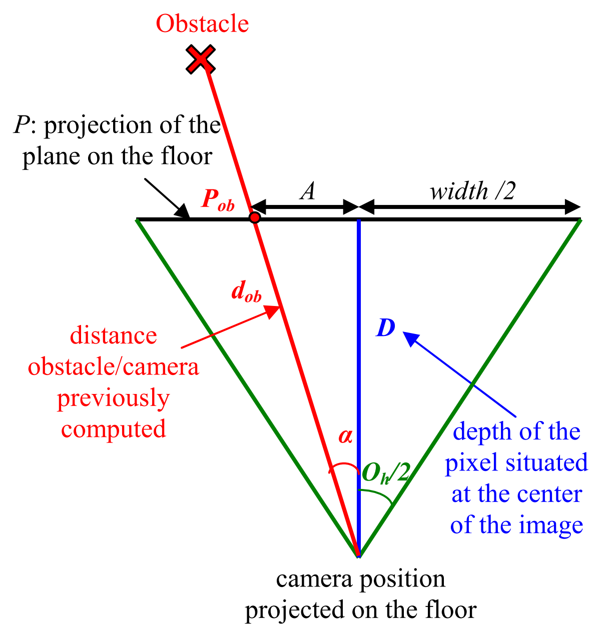

Fig.9: Obstacle coordinates calculation

IV. Experimental Results

In order to develop, test and validate our applications, methods and algorithms, real sensor data are used. A sensor data acquisition platform was then developed. This platform is developed and integrated on our experimental vehicle *CyCab* developed by the Robosoft[2] company, and can acquire, date and record measurements of our on-board sensors.

[2] http://www.robosoft.com/eng/

A. *Experimental results of the obstacle detection method*

Outdoor experiments have been realized in order to validate our obstacle detection and tracking method. The experimental protocol of obtained results is realized as follows. The GPS and the camera are installed in our experimental vehicle *CyCab*. The on-board camera is then calibrated, and calibration results are used by the 3D-GIS to define the parameters of the virtual camera. The experimental area is the Stanislas square at Nancy, in France. A non-predefined trajectory is realized on this square. The developed graphical interface, which presents results, is Figure 10.

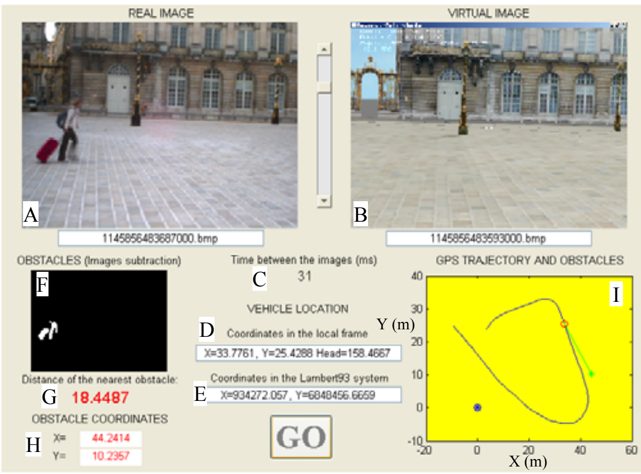

Fig.10: Results display

Legends from *A* to *I* are used in the follow to describe each part of this interface:

- Part A: Real image acquired by the on-board camera and its acquisition date in nanoseconds (time since 1970/01/01) at the bottom of the image
- Part B: Virtual image computed by the 3D-GIS and its acquisition date
- Part C: Time difference between the 2 dates of acquisition in milliseconds
- Part D: Display of the GPS pose in the 3D model local coordinate system
- Part E: GPS pose in the *Lambert93* frame
- Part F: Image processing result: image containing pixels representing the detected obstacle(s)
- Part G: Distance between the vehicle and the detected obstacle(s)
- Part H: Calculated geo-position of the obstacle(s) in the local frame
- Part I:
  - The continuous blue line is test path i.e. plot of the centimetric LRK GPS position converted in the local model frame. The origin of this local frame, represented by the blue circled star, is Stanislas statue
  - The red circle on the test path is the current position of the vehicle
  - Part I background become yellow if an obstacle is detected (alarm)
  - The green star is the obstacle
  - The green segment is the link between the vehicle and the obstacle, i.e. the estimation of the distance

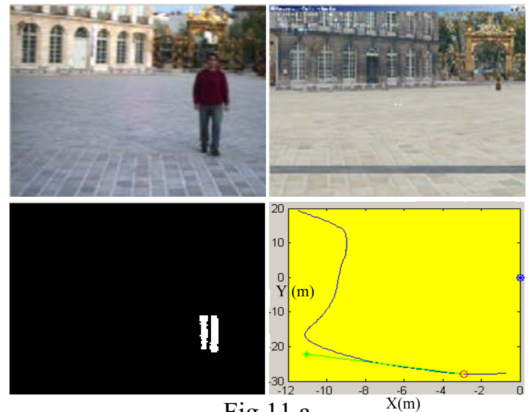

Fig.11.a

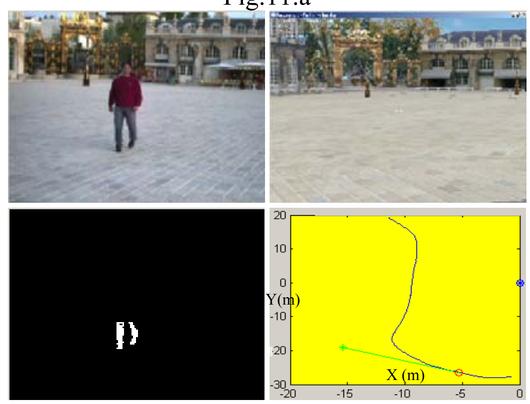

Fig.11.b

Fig.11: Obstacles detection results

In Figure 11, two others results of the presented obstacle detection application are shown. For each example and to simplify, just four elements are displayed: the real image (top-left), the virtual image (top-right), the substraction image, which represents the obstacles (bottom-left) and the plot Figure (bottom-right) with the complete trajectory (centimetric LRK GPS position), the current position (red circle), the obstacle (green star) and the alarm (yellow color background) in case of detected obstacle. During the carried out experiments, the smallest detected obstacle was a pigeon and the biggest one was a person.

B. *Results validation with Ladar*

The distance between the vehicle and the obstacle has been computed with our vision/3D model method. In order to validate this distance, we compare our results with ladar measurements. To use both the ladar and the camera in the same framework, the geometric transformation between the telemeter and the camera references has to be known.

Usually for mobile perception sensors, we make the distinction between intrinsic and extrinsic models. The intrinsic model gives the relationship between the observed reality and its measures, in a frame relative to the sensor, whereas the extrinsic model gives the transformation between the relative frames of the sensor to a fixed frame.

This distinction applies to both the camera and the laser range scanner. As the two sensors are rigidly fixed on a common rigid support, the position and orientation between the telemeter and the camera can be manually and rigorously approximate. A GPS antenna mounted on the top of the support permits to geo-reference the commun support.

Fig.12 presents the vehicle/obstacle distance comparison for a person that walks in front of the moving vehicle during 40 seconds. In Fig.12.a, the distance between the vehicle and the obstacle is plotted in dark (blue) for our vision/3D model method and in bright (yellow) for ladar measurement. In Fig.12.b, the difference between those two distance values is plotted. One can remark that the maximal error is 0.8 meter.

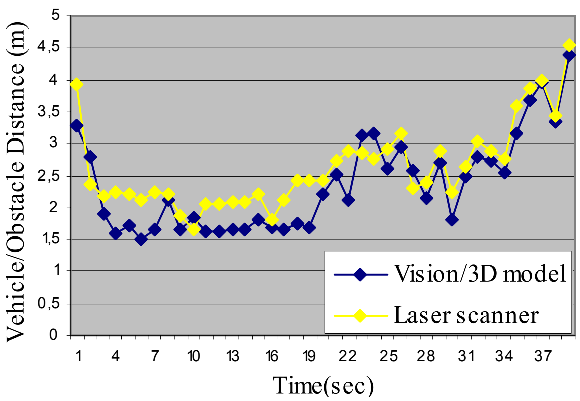

Fig.12.a

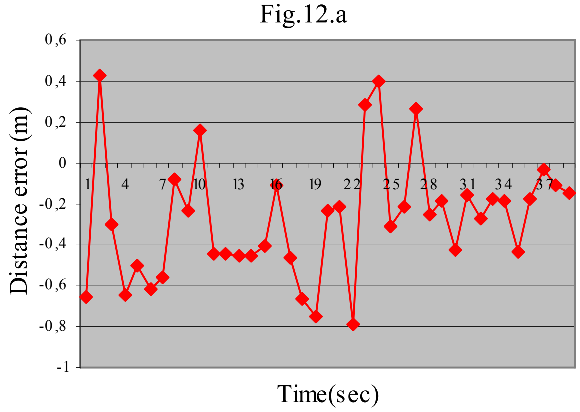

Fig.12.b

Fig.12: Vehicle/Obstacle distance comparison

V. Conclusion

An obstacle detection and geo-localization method has been developed. This method uses a GPS, a monovision camera and a geographical 3D model managed by a 3D-GIS, which is a new kind of information source. The obtained distance was compared and validated with laser scanner measurements. A more robust and precise observation of the vehicle pose should still improve the results. GPS measurements can, for example, be fused with other sensors (odomter, compass …). The ladar/camera calibration process has also to be performed. Moreover, the method described to compute the geo-localization of the obstacle had to be extended in order to take into account that floor ground planes can not be guaranteed in the real world.

References

[1] R. Aufrère, F. Marmoiton, R. Chapuis, J. P. Derutin, and F. Collange, “Road detection and vehicle tracking by vision for Adaptative Cruise Control”, *International Journal of Robotics Research*, vol. 20, n°4, pp. 267-286, April 2001.

[2] A. Bensrhair, M. Bertozzi, A. Broggi, P. Miche, S. Mousset and G. Toulminet, “A cooperative approach to vision-based vehicle detection”, *in Proc. Intelligent Transportation System Conference*, Japan, October 2001.

[3] M. Betke, E. Haritaoglu and L. Davis, “Multiple vehicle detection and tracking in hard real-time”, *in Proc. IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, Sept. 1996, pp. 351–356.

[4] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski and F. Charpillet, “Localisation in urban environment using GPS and INS aided by monocular vision system and 3D geographical model”, *in Proc. IEEE Intelligent Vehicles Symposium*, IV'07, Istanbul, Turquie, 13 - 15 Juin, 2007.

[5] U. Franke and A. Joos, “Real-time stereo vision for urban traffic scene understanding”, *in Proc. IEEE Intelligent Vehicle Symposium*, Dearborn, USA, October 2000.

[6] M. Gavrila, M. Kunert and U. Lages, “A multi-sensor approach for the protection of vulnerable traffic participants - the protector project”, *in Proc. IEEE Instrumentation and Measurement Technology Conference*, Budapest, Hongrie, May 2001.

[7] J. Hancock, M. Hebert and C. Thorpe, “Laser intensity-based obstacle detection”, *in Proc. IEEE Conference on Intelligent Robots and Systems*, IROS’98, October 1998.

[8] C. Knoeppel, A. Schanz and B. Michaelis, “Robust vehicle detection at large distance using low resolution camera”, *in Proc. IEEE Intelligent Vehicle Symposium*, Dearborn, USA, October 2000.

[9] D. Koller, K. Daniilidis, and H. Nagel, “Model-based object tracking in monocular image sequences of road traffic scene”, *International Journal of Computer Vision*, vol. 3, n°10, pp. 257-281, 1993.

[10] Q.T. Luong, J.Weber, D. Koller and J. Malik, “An integrated stereo-based approach to automatic vehicle guidance”, *in Proc. 5th International Conference on Computer Vision*, pp. 52-57, Cambridge, MA, USA, June 1995.

[11] R. Malaka, K. Schneider, U. Kretschmer, “Stage-based augmented edutainment”, *in Proc. Smart Graphics: 4th International Symposium*, SG 2004, Banff, Canada, May 23-25, 2004.

[12] M. Parent and M. Crisostomo, “Collision avoidance for automated urban vehicles”, *in Proc. IEEE Intelligent Vehicles Symposium 2001*, IV 2001, Tokyo, Japan, June 2001.

[13] S. Persa and P. Jonker, “Real-time Computer Vision System for Mobile Robot”, *in Proc. SPIE Intelligent Robots and Computer Vision XX : Algorithms, Techniques, and Active Vision, David P. Casasent, Ernest L. Hall, Eds*., vol. 4572, pp. 105-114, October 2001.

[14] K. Sung-Soo, K. Kyong-Ho, L. Seong-Ho and L. Jong-Hun, “High Resolution Earth Imaging for Geospatial Information”, *Video Navigation System using the Geographic Hypermedia*, ISPRS Hannover Workshop, 2005.

[15] Y. Takase, N. Sho, A. Sone, K. Shimiya, “Generation of Digital City Model”, *Journal of the Vizualization Society of Japan*, vol. 23, n°88, pp. 21-27, 2003.

[16] J.C. Van den Heuvel, J.C.M. Kleijweg, W. van der Mark, C.M. Lievers and L.J.H.M. Kester, “Obstacle detection for people movers using vision and radar”, *in Proc. 10th World Conference on Intelligent Transport Systems and Services*, Madrid, Spain, 16-20 November 2003.

[17] T. Williamson and C. Thorpe, “Detection of small obstacles at long range using multibaseline stereo”, *in Proc. IEEE Intelligent Vehicles Symposium*, Stuttgart, Germany, Oct. 1998.

[18] S. Zlatanova, A. Rahman and M.Pilouk, “Present status of 3D GIS”, *GIM International*, vol. 16, n°6, pp. 41-43, June 2002.

[19] Z. Zhang, “A flexible new technique for camera calibration”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, N°11, pp 1330-1334, 2000.

[20] Calibration Camera Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html