Free Space Estimation for Autonomous Navigation

Nicolas Soquet[1], Mathias Perrollaz[1], Raphael Labayrade, Didier Aubert

LIVIC (INRETS / LCPC)
Bât. 824, 14, route de la Minière
78000 Versailles - France
nicolas.soquet@lcpc.fr mathias.perrollaz@inrets.fr

Abstract. One of the issues in autonomous navigation is the free space estimation. This paper presents an original framework and a method for the extraction of such an area by using a stereovision system. The *v-disparity* algorithm [11] is extended to provide a reliable and precise road profile on all types of roads. The free space is estimated by classifying the pixels of the disparity map. This classification is performed by using the road profile and the *u-disparity* image. Each stage of the algorithm is presented and experimental results are shown.

1 Introduction

One of the issues in autonomous navigation is the free space estimation. Many systems have been designed to deal with obstacle detection in various environments. Many methods were proposed using optical flow [4] [7], radars [8] [15], laser range finder [9] [14], stereovision [16] [11] and multi-sensor fusion on structured roads. The 2005 DARPA Grand Challenge competition was successfully achieved on unstructured roads [5]. Most of the teams used lidar, radar or sonar [1] but no one succeed with a single vision-based system. Such systems have strong advantages such as low cost (critical issue for industrial matter) and passivity (required for military systems). The TerraMax team completed the race with a stereovision-based obstacle detection [3] and the help of other active sensors [13]. This shows the potential use of stereovision. In this paper we present a system based on stereovision for both structured and unstructured roads. We use a robust method to extract a precise profile of the road surface which allows us to estimate the free space. The paper is organized as follows: section 2 presents the stereoscopic sensor and the set of equations used. Section 3 proposes adaptations of the *v-disparity* algorithm for the unstructured roads. Section 4 explains how we collect the road profile. Section 5 shows how the free space is estimated. Finally section 6 presents experimental results obtained from real images.

2 Geometrical aspects

Our system uses a single stereovision sensor (Fig. 1) which will be described in section 2.1. The transformations between the disparity space (linked to the

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

images) and the real space (linked to the vehicle) are described in the section 2.2.

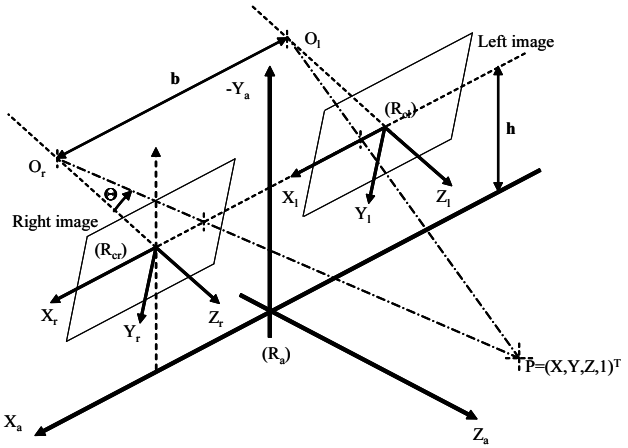

Fig. 1. Geometry of the stereoscopic sensor and coordinate systems (a: absolute, r: right, l:left).

2.1 Stereoscopic sensor

The stereovision sensor (Fig. 1) uses two cameras positioned at the same height relative to the ground level. An adjustment is performed so that the epipolar lines are parallel and correspond to the scanning lines. The parameters described on the figure are:

- h: height of the cameras above the ground,
- b: stereo baseline (distance between the cameras),
- θ: pitch of the cameras (angle between the horizontal and the optical axis).

The coordinates (u, v) give the position of a point in an image plane. The intrinsic parameters of the cameras are the focal length of the lens (f) and the size of the pixels (t_u, t_v). According to the used cameras specifications, we can approximate: $t_u \approx t_v = t$. Then we use $\alpha = f/t$.

2.2 Coordinate systems

(R_a) is the absolute coordinate system. (R_r) is the coordinate system of the right image and (R_l) is the one of the left image. For a given point, the R_a-coordinates are denoted by (X, Y, Z), the R_r-coordinates by (u_r, v_r) and the R_l-coordinates by (u_l, v_l). Δ is the disparity value and is equal to $u_l - u_r$. Thanks to the epipolar

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

constraint we have $v_r = v_l$. We call disparity space the (u_r, v, Δ) space.
Using the pin-hole camera model and the projection of the optical center (u_0, v_0), we obtain the following equations (1):

$$\left|\begin{array}{ll} u_l & = u_0 + \frac{\alpha X + \alpha b/2}{(Y+h)\sin\theta + Z\cos\theta} \\ u_r & = u_0 + \frac{\alpha X - \alpha b/2}{(Y+h)\sin\theta + Z\cos\theta} \\ v_r = v_l & = \frac{[v_0\sin\theta + \alpha\cos\theta](Y+h) + [v_0\cos\theta - \alpha\sin\theta]Z}{(Y+h)\sin\theta + Z\cos\theta} \\ \Delta & = u_l - u_r = \frac{\alpha b}{(Y+h)\sin\theta + Z\cos\theta} \end{array}\right. \quad (1)$$

Those equations (1) describe the transformation between the image coordinate systems (R_r and R_l) and the real space (R_a).

2.3 Plane projection

The scene is often modelled by a set of planes in the real space. It is proven in [11] that a 3D plane estimation can be reduced to a 2D straight line estimation by using a well defined plane projection. Therefore if the scene is simply modelled with a planar road and vertical obstacles, then finding the good plane projection is sufficient to easily estimate the planes. It is shown in [11] that the *v-disparity* image is very efficient for such estimations.

The v-disparity image is a plane projection. $\boldsymbol{u}$ is the normal and $(\boldsymbol{v}, \boldsymbol{\Delta})$ is the basis of the plane (denoted P_v). The *v-disparity* image is computed by accumulating the pixels of same disparity along the $u-$axis. Let consider a pixel (denoted P) on the *v-disparity* image with coordinates (v_p, δ_p). The intensity of P equals the number of pixels on the line v_p of the disparity map, having a disparity of δ_p. Thanks to those accumulations the *v-disparity* image is very robust with regard to the noise. This image can be used to extract the road plane and vertical obstacles (see [11]).

The u-disparity image is a plane projection. $\boldsymbol{v}$ is the normal and $(\boldsymbol{u}, \boldsymbol{\Delta})$ is the basis of the plane. The *u-disparity* image is computed by accumulating the pixels of same disparity along the $v-$axis. Like the *v-disparity* image, the *u-disparity* image tolerates the noise due to the accumulation process. However it is not used to extract any plane. It is used to detect obstacle pixels. Indeed high intensity in the *u-disparity* image indicates that many pixels in the disparity map have the same disparity in a single column, thus they do not belong to the road. They correspond to vertical alignments in the real space and can be considered as obstacles. [2] uses a similar approach for obstacle detection by marking as obstacles the regions with similar disparity.

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

3 An enhanced v-disparity algorithm

3.1 Basic algorithm for a global road profile

The computation of the *v-disparity* image is fully described in [11]. The algorithm was designed to work on structured roads. First the primitives are computed using horizontal local maxima gradient on both images and a single threshold. Then the disparity map is computed by local matching, based on normalized correlation around the local maxima with a left/right checking failure procedure [6]. The *v-disparity* image is computed from the disparity map by accumulating the pixels of same disparity along the $v-$axis. A global road profile can be easily extracted from the *v-disparity* image using any straight line extracting procedure, such as the Hough transform (see Fig. 2).

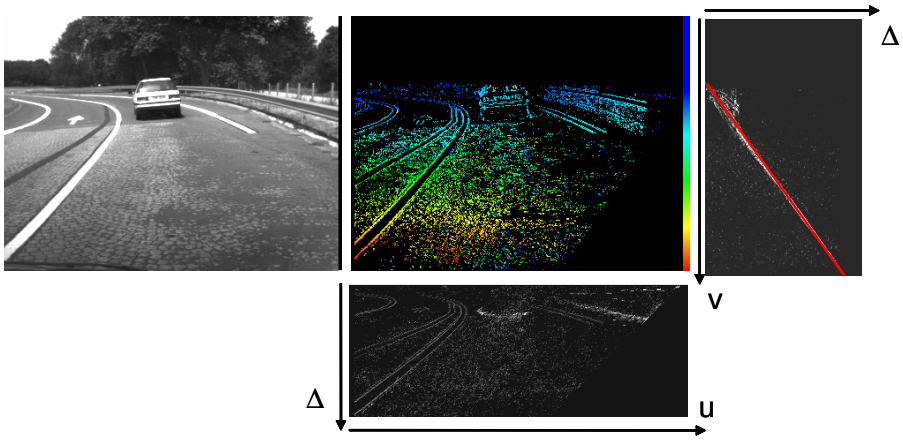

Fig. 2. Top: left to right: an image of the stereo pair corresponding to a structured road, the disparity map and the *v-disparity* image with the global road profile colored in red. Bottom: the corresponding *u-disparity* image. The images were computed with a gradient threshold of 6.

3.2 Adaptive thresholds

When we have tested this basic algorithm on unstructured roads, we have extracted wrong global road profile in some cases. It happened because the primitives were badly distributed in the images. The Hough transform did not take the best alignment in the *v-disparity* image. An alignment of few points with high intensity were preferred to an alignment of a larger number of points with lower intensity. On Fig. 3, the global road profile extracted is wrong. In this example, the numerous pixels corresponding to the trees on the horizon line accumulate to

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

form a vertical alignment in the *v-disparity* image. Its intensity is much higher than the alignment corresponding to the road profile. With structured roads we can find a threshold to select primitives which give good results. However it is harder to find such a threshold in unstructured roads.
We use adaptive thresholds to deal with this issue. Instead of setting a common threshold for both images, we set a percentage of primitives per line. This percentage is the same for all the lines of both images. Then we select a threshold for each line using a cumulative histogram (see Fig. 3). This algorithm is efficient to get a good primitives distribution. The Fig. 3 shows the results on the *v-disparity* and the global road profile. With a constant threshold of 6, the global road profile is wrong. With a constant one of 2, the profile is correct, but there is a lot of noise. With the adaptive thresholds, the profile is correct and the *v-disparity* image contains less noise.

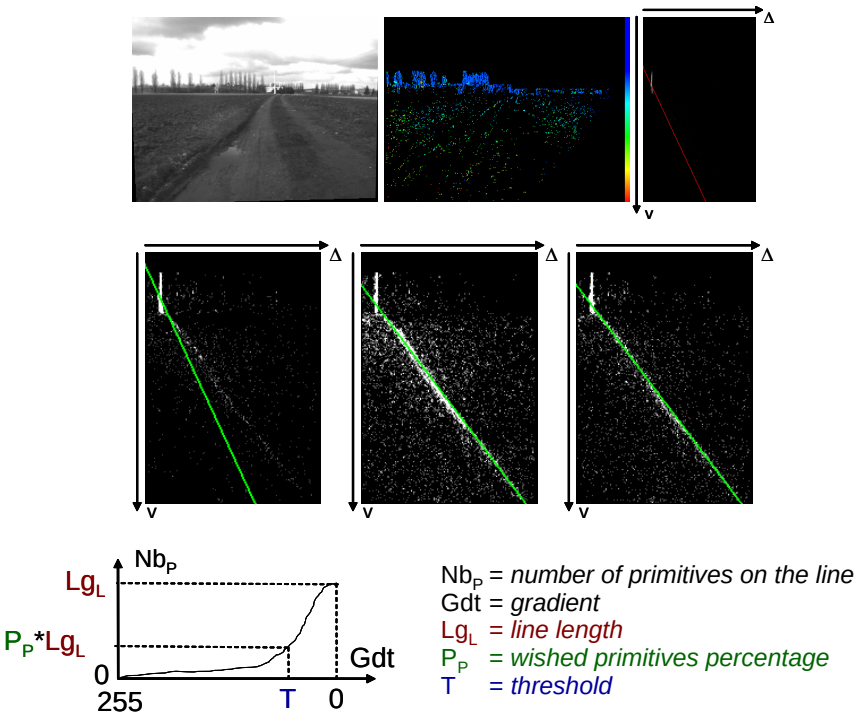

Fig. 3. *Top*: left to right: an image of the stereo pair corresponding to a unstructured road, the disparity map and the *v-disparity* image with the global road profile in red. The images were computed with a gradient threshold of 6. *Middle:* left to right: *v-disparity* image of a non structured road and the global profile in green for a threshold of 6, 2 and adaptive ones. *Bottom:* illustration of the cumulative histogram.

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

3.3 Column normalization of the v-disparity image

Most of the time the adaptive thresholds provide a correct global road profile. However in some cases a bad distribution still disturbs the Hough transform. Indeed the adaptive thresholds can not guarantee low intensity variations in the *v-disparity* image. High intensity pixels can still disturb the extraction of straight lines. In some particular cases such as a large obscuring obstacle in the image, the Hough transform may not extract an alignment corresponding to the road profile, but a wrong one due to the high intensity of the obstacle pixels. When an obstacle fits in a vertical alignment of the *v-disparity* image, we process a normalization on each column of the *v-disparity* image. The intensity of each pixel is divided by the highest intensity in its own column. This way the influence of the pixels corresponding to obstacles is decreased and the longest alignment is more relevant. Since the longest alignment in the *v-disparity* image suits the road profile, the Hough transform is more reliable.

4 Road profile

4.1 Global profile

The enhanced algorithm provides an efficient way to extract the global road profile. Such a profile is not suitable for unstructured roads, because one straight line in the *v-disparity* image corresponds to a planar road. We need a better accuracy (roll, yaw, pitch or road slant) with unstructured environments. Some techniques were designed to solve this. We can segment the *v-disparity* image along the v-axis and extract a straight line for each bins to deal with the altitude variation along the road (uphill or downhill). Thus the road is modelled as a succession of planes. The roll, pitch and yaw can be estimated by an extension of the *v-disparity* algorithm to other planes extraction [10] or a local analysis of the *v-disparity* [12].

4.2 Precise profile

Our approach to obtain a more precise profile of the road is original and different to the previously described ones [10] [12]. We exploit the *v-disparity* image to the maximum. The goal is to detect all the pixels belonging to the road on the *v-disparity* image. Thus we extract the exact shape of the road profile in the *v-disparity* image using a propagation algorithm. The algorithm consists of two different phases: an initialization one and a propagation one.

The initialization phase consists in extracting the global profile as previously described. Then the pixels on the *v-disparity* image belonging to the global profile are selected. Only the pixels with high intensity value are kept. Those pixels are the seeds for the next propagation phase. This method preserves the robustness of the *v-disparity* representation, because it keeps the most cumulating pixels and removes the noisy ones.

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

The propagation phase needs two parameters: a threshold on the intensity of the *v-disparity* pixels and a specific pattern to give the direction of the propagation. We start the propagation from the seed pixels. For each tested pixel we check if the intensity value is high enough (above the specified threshold). If it is the case, we add this pixel (denoted P) to the precise profile and we test all the pixels related to P, according to the used pattern centered on P. Otherwise we do nothing, P is left and we restart from another seed. We go on until each seed has been propagated.

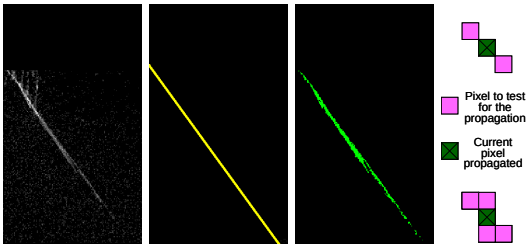

Fig. 4. *Left to right: v-disparity* image, the global profile in yellow color, the precise one in green color and two examples of possible patterns. *Concerning the patterns:* the bottom one is unsuitable because it allows propagation in the column of the current pixel. Indeed it may propagate an obstacle which must not belong to the precise profile. The top one is simple and appropriate because it allows a propagation in directions corresponding to the orientation of the profile.

5 Free space estimation

The precise profile is then used to classify the pixels: the ones belonging to the road surface and the others supposed to be obstacles or noise. To perform this classification the *u-disparity* image is used as well.

5.1 Pixel classification

First the classification phase does not process all the pixels of the image, because the disparity value is required. Thus it processes only the pixel of the disparity map which have a non-zero disparity value. That means that only the primitives, computed with the adaptive thresholds (see 3.2), which are matched, will be classified. Therefore for such a pixel (denoted P) the non-zero disparity value gives a projection point in the *u-disparity* image (denoted P_u) and a projection point in the *v-disparity* image (denoted P_v). Then the classification algorithm works as follow: if the intensity value in the *u-disparity* image of P_u is high then P is classified as an obstacle pixel. Otherwise if P_v belongs to the precise profile P is classified as a road surface pixel, otherwise P is not classified (it is supposed

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

to be noise). This classification phase does not consider all pixels. However the ones classified are very reliable due to the robustness of the *v-disparity* and *u-disparity* images. That is why they are used to extend the classification to the unclassified ones with a propagation phase.

5.2 Propagation of the free obstacle area

The propagation of the free obstacle area is straightforward and based on the robustness of the classification. The reliable pixels are classified into road surface and obstacles. An influence area is computed for each of them by using a two dimensional gaussian surface. Thus all pixels receive a contribution from all the classified pixel. The value of each contribution depends on the distance to the classified pixel and its class. If it is a road surface pixel the value is positive. Otherwise it is negative. The longer is the distance, the lower is the absolute value. Then all pixels cumulate the value of the contribution of all classified pixels. Afterwards an unclassified pixel surrounded by surface road pixels has a positive sum and is labelled as road surface. On the other hand an unclassified pixel surrounded by obstacle pixels has a negative sum and is labelled as obstacles. In this manner all unclassified pixels are labelled. This phase is not as robust as the classification one, but it is a good way to show how the classification using a precise profile is efficient. In order to increase the robustness we plan to add a checking procedure of the final propagation.

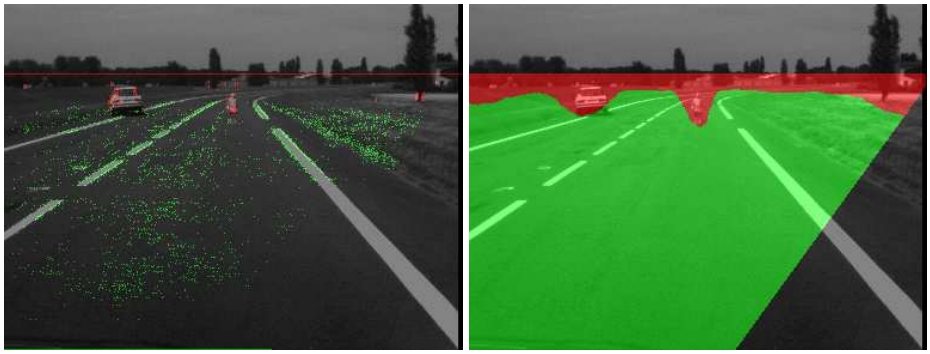

Fig. 5. Left: the result of the classification step, the green points correspond to the road classified pixels and the red ones to the obstacle classified pixels. Right: the result of the propagation one: the green area is the free space.

6 Results

The method presented in this paper has been implemented and tested with real video sequences on a 2.4 GHz computer. Figure 6 shows some results obtained

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

in various environments. In most cases the free space is well estimated. The classification phase gives good results, but when too few pixels are classified, the propagation phase can fail. The number of primitives can be increased with the adaptive thresholds to obtain a denser disparity image. This would increase the computation load and prevent real time processing.

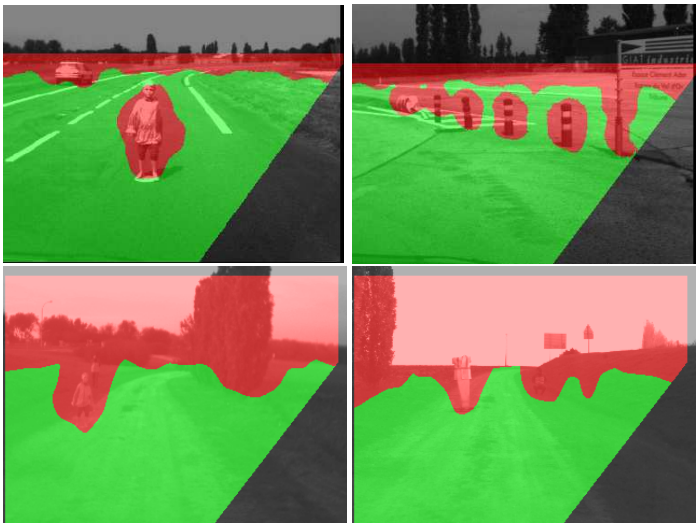

Fig. 6. Translucent green (or light translucent grey value): the free space estimation, translucent red (or dark translucent grey value): the obstacle areas estimation.

7 Conclusion

This paper presented an original method to estimate the free space by using a stereovision system. It uses the robustness of the *u-disparity* and *v-disparity* representations. First the original *v-disparity* algorithm is extended to unstructured roads. Adaptive thresholds and column normalization of the *v-disparity* image enable the extraction of a good global road profile. The precise profile takes into account all road variations (yaw, pitch, roll, bumpy, damaged) and allows to work on any types of roads. A classification process is performed to classify the pixels of the disparity map as road surface or obstacle. This phase gives very good results. Finally the propagation of the free obstacle area labels all the remaining pixels. This last phase can be improved by using morphological processing (opening and closing operations) to remove small artefacts or to better emphasize the obstacles. Other developments are considered. Working on the disparity map density and reducing the computation time are key improvements. Currently we are studying a procedure to check the free space propagation.

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

8 Acknowledgments

The authors would like to thank Joel G. Morillon and Cecile Fialaire from Thales company and Dr. Michel Devy from LAAS for their advices and contributions. This work was funded by the French Defense Procurement Agency.

References

1. R. Behringer, S. Sundareswaran, B. Gregory, R. Elsley, B. Addison, W. Guthmiller, R. Daily, and D. Bevly. The darpa grand challenge - development of an autonomous vehicle. In *IEEE Intelligent Vehicle Symposium (IV'04)*, Parma, Italy, June 2004.
2. A. Broggi, C. Caraffi, R. I. Fedriga, and P. Grisleri. Obstacle detection with stereo vision for off-road vehicle navigation. In *IEEE Machine Vision for Intelligent Vehicles*, San Diego, USA, June 2005.
3. A. Broggi, C. Caraffi, P. P. Porta, and P. Zani. The single frame stereo vision for reliable obstacle detection used during the 2005 darpa grand challenge on terramax. In *IEEE Intelligent Transportation Systems*, Toronto, Canada, Sept. 2006.
4. S. Carlsson and J.-O. Eklundh. Object detection using model based prediction and motion parallax. In *Proc. of the Workshop on Vision (Prometheus, Pro-Art)*, volume 427, pages 297–306, Sophia Antipolis, France, 1990.
5. DARPA. Grand challenge 2005. [online]: http://www.grandchallenge.org.
6. G. Egnal and R. P. Wildes. Detecting binocular half-occlusions: Empirical comparisons of five approaches. *IEEE Transactions on pattern analysis and machine intelligence*, 24(8):1127–1133, 2002.
7. W. Enkelmann. Object detection by evaluation of optical fields. In *Proc. of the Workshop on Vision (Prometheus, Pro-Art)*, volume 427, pages 146–166, Sophia Antipolis, France, 1990.
8. T. Kato, T. Tanizaki, T. Ishii, H. Tanaka, and Y. Takimoto. 76 ghz high performance radar sensor featuring fi ne step scanning mechanism utilizing nrd technology. In *IEEE Intelligent Vehicles Symp. (IV'01)*, Tokyo, Japan, June 2001.
9. A. Kirchner and A. Ameling. Integrated obstacle and road traking using a laser scanner. In *IEEE Intelligent Vehicles Symp. (IV'00)*, Dearborn, USA, Oct. 2000.
10. R. Labayrade and D. Aubert. A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision. In *IEEE Intelligent Vehicle Symposium (IV'03)*, Columbus, USA, June 2003.
11. R. Labayrade, D. Aubert, and J. Tarel. Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation. In *IEEE Intelligent Vehicle Symposium (IV'02)*, Versailles, France, June 2002.
12. V. Lemonde and M. Devy. Obstacle detection with stereovision. In *IEEE Mechatronics and Robotics (MECHROB'04)*, Aachen, Germany, September 2004.
13. U. Ozguner, K. A. Redmill, and A. Broggi. Team terramax and the darpa grand challenge: A general overview. In *IEEE Intelligent Vehicle Symposium (IV'04)*, pages 232–237, Parma, Italy, June 2004.
14. M. Parent and M. Crisostomo. Collision avoidance for automated urban vehicles. In *IEEE Intelligent Vehicles Symposium (IV'01)*, Tokyo, Japan, June 2001.
15. T. Uebo, T. Kitagawa, and T. Iritani. Short range radar utilizing standing wave of mircowave or millimeter wave. In *IEEE Intelligent Vehicles Symposium (IV'01)*, Tokyo, Japan, June 2001.
16. T.A. Williamson. A high performance stereo vision system for obstacle detection. In *PhD thesis*, Carnegie Mellon University, 1998.