ONBOARD ROAD OBSTACLES DETECTION IN NIGHT CONDITION USING BINOCULAR CCD CAMERAS

Raphael Labayrade
Didier Aubert
Sio Song Ieng
Vehicle-Infrastructure-Driver Interactions Research Unit
LIVIC - INRETS/LCPC
FRANCE
Paper Number 109

ABSTRACT

This paper presents a CCD based stereovision algorithm (called the ”v-disparity” algorithm [1]) for onboard road obstacles detection (pedestrians, vehicles, motorbikes, etc.) in night condition. The algorithm is explained, and then evaluated towards different obstacles. The theoretical good properties of the ”v-disparity” algorithm - accuracy, robustness, computational speed - are experimentally confirmed. Experiments show that obstacles are detected in a precise manner with high confidence values, at frame rate (25 Hz) using no special hardware.

1 INTRODUCTION

In the context of Intelligent Transportation Sytems, onboard road obstacles detection is an essential task, in both day and night condition. Various sensors can be used in this purpose. Radar or Lidar constitute one possibility but present some drawbacks : they are expensive and delicate devices, and send electromagnetic waves into the environment. For night vision, infrared cameras can be used but remain expensive. CCD based stereovision is a cheaper passive sensor for detecting road obstacles, but most stereovision based algorithms are designed for detecting obstacles in day condition. This paper presents a stereovision based algorithm (called the ”v-disparity” algorithm [1]) that can detect generic obstacles (pedestrians, vehicles, motorbikes, etc.) in both day [2] and night condition, using two CCD cameras. Experimental results of detection in night condition are presented : obstacles are detected in a robust and precise manner in the area located under the lighting of the equiped vehicle (up to 40 meters from the vehicle). The paper stresses the theoretical and experimental properties of the algorithm that allow its efficient use in the automotive context where lighting conditions are difficult. The paper is organized as follows.

Section 2 presents our experimental protocol. Section 3 deals with the description of our algorithm (called the ”v-disparity” algorithm), and the method used for robustly detecting generic obstacles. A disparity map is computed and a geometric representation of the road scene is deduced in an original manner. Section 4 presents experiments that evaluate the efficiency of the ”v-disparity” algorithm for detecting different obstacles in night condition. All tested obstacles are detected with a good confidence value. Section 5 evaluates the robustness of the method against difficult meteorologic conditions (rain in night condition) and against noise. Computing time is also evalutated : the detection process is performed within 40 ms using a 733 MHz Pentium III processor. Eventually, Section 6 deals with future work.

2 EXPERIMENTAL PROTOCOL

Fig. 1 presents the stereo sensor used for the experiments (top), the left CCD camera (bottom left), and the configuration area used for the evaluation of the obstacle distance accuracy (bottom right), shown in day light. After configuration, the image planes are parrallel : the epipolar geometry is rectified (epipolar lines correspond to scanning lines in the images of the stereo pair). The parameters of the stereo sensor are $b = 1.03$ m, $h = 1.4$ m, $\theta = 11.3°$, $f = 8.5$ mm, $t_u \approx t_v = 7.2\ \mu m$. The resolution of each image is 380×288 pixels ($\frac{1}{4}PAL$). $Computar^{TM}$ auto-iris lenses and a $Matrox^{TM}$ Meteor II board are used for grabbing images on a $PIII$ 733 GHz PC computer running under Microsoft Windows 2000 ©. With this configuration the disparity range investigated is $[0, 150]$ pixels.

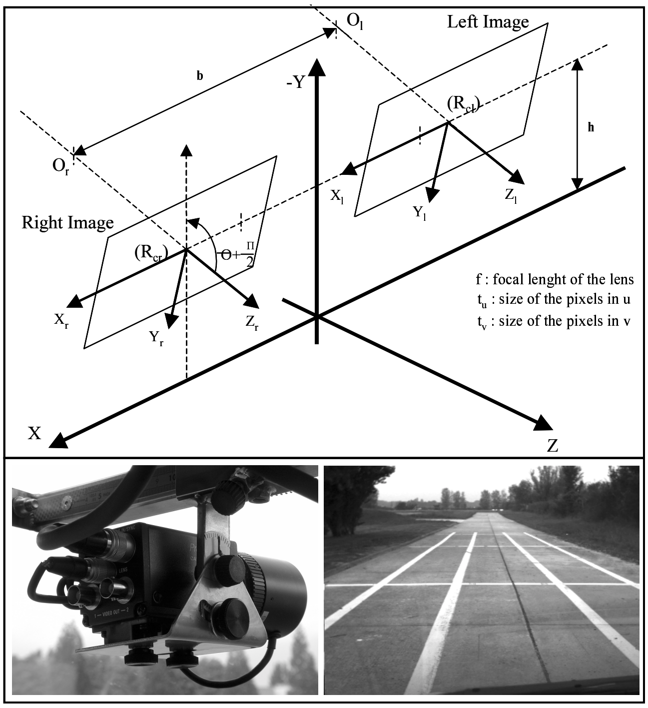

Figure 1: **Stereo sensor (top). Left camera (bottom left). Configuration area (bottom right)**.

3 THE ”V-DISPARITY” ALGORITHM

3.1 Algorithm overview

A framework for our obstacles detection process is presented in Fig. 2 (left). Our current implementation of this framework is as follows (see Fig. 2 (right)): first, a pair of stereoscopic images is grabbed. A sparse disparity map is then computed. The ”v-disparity” image is build and global surfaces are extracted. The position of obstacles on the road surface are then deduced. This algorithm is detailed in the next part of this section.

3.2 Modeling of the stereo sensor

The two image planes of the stereo sensor are supposed to belong merely to the same plane and are at the same height above the road (see Fig. 1). This camera geometry

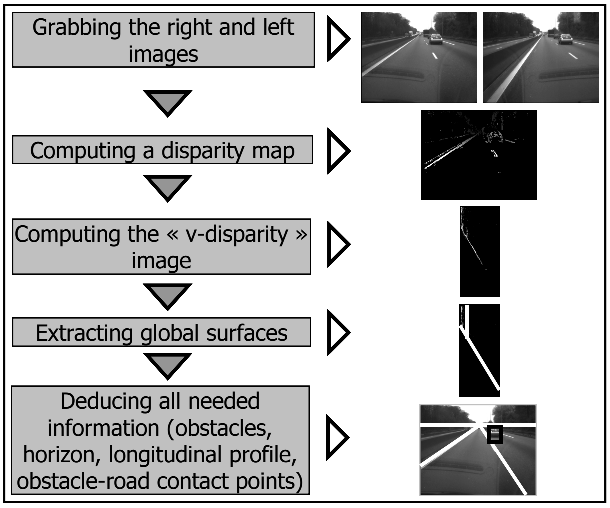

Figure 2: **Framework (left) and exemple of implementation (right) of the ”v-disparity” algorithm.**

means that the epipolar lines are parallel.

In what follows we will need to perform positioning in three coordinate systems shown in Fig. 1: R_a (absolute), R_{cr} (right camera) and R_{cl} (left camera). R_a is the road coordinate system.

The other parameters on the figure are as follows:

- θ : is the angle between the optical axis of the cameras and the horizontal,
- h : is the height of the cameras above the ground,
- b : is the distance between the cameras (i.e. the stereoscopic base).

In the camera coordinate system, the position of a point in the image plane is given by its coordinates (u, v). The image coordinates of the projection of the optical center will be denoted by (u_0, v_0), assumed to be at the center of the image. The intrinsic parameters of the camera are f (the focal length of the lens), t_u and t_v (the size of pixels in u and v). We also use $\alpha_u = f/t_u$ and $\alpha_v = f/t_v$. With the cameras in current use we can make the following approximation: $\alpha_u \approx \alpha_v = \alpha$.

Using the pin-hole camera model, a projection on the image plane of a point (X,Y,Z) in R_a is expressed by:

$$\begin{cases} u = \alpha_u \frac{X}{Z} + u_0 \\ \\ v = \alpha_v \frac{Y}{Z} + v_0 \end{cases} \tag{1}$$

On the basis of Fig. 1, the transformation from the absolute coordinate system to the camera coordinate system is achieved by the combination of a vector translation $\vec{t} = -h\vec{Y} + \varepsilon_i \frac{b}{2}\vec{X}$ (with $\varepsilon_i = -1$ in R_{cl} or 1 in R_{cr}), and a rotation around $\vec{X}$ by an angle of $-\theta$. Let T_i denote the translation matrix, R the rotation matrix and $D_i = RT_i$. In homogeneous coordinates, the different transformation matrices are therefore:

$$T_i = \begin{pmatrix} 1 & 0 & 0 & -\varepsilon_i \frac{b}{2} \\ 0 & 1 & 0 & h \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2}$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3}$$

where i is equal to r, l (right or left).

It is necessary to perform a perspective projection in order to express fully the coordinates of the points in the image plane coordinate system. The perspective projection matrix M_{proj} is expressed as follows:

$$M_{proj} = \begin{pmatrix} \alpha_u & 0 & u_o & 0 \\ 0 & \alpha_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{4}$$

Finally, we obtain the matrix of transformation T_{ri} from the absolute coordinate system R_a to the image coordinate system i (i is equal to l or r):

$$T_{ri} = M_{proj} D_i \tag{5}$$

If P is a point with homogeneous coordinates $(X, Y, Z, 1)^T$ in R_a, its homogeneous coordinates in the image coordinate system i are:

$$p = T_{ri} P = (x, y, z)^T \tag{6}$$

We can then compute the non-homogeneous (u, v) coordinates of P as:

$$\begin{cases} u = \frac{x}{z} \\ \\ v = \frac{y}{z} \end{cases} \tag{7}$$

3.3 Modeling of the road

In what follows we will consider that the road is modeled as an horizontal plane with equation $Y = 0$.

3.4 Modeling of the obstacles

In what follows we will consider that any obstacle is characterised by a vertivcal plane with equation $Z = d$.

Thus, all planes of interest can be characterised by a single equation: $Z = aY + d$.

3.5 The image of planes of interest in the "v-disparity" image

Let P be a point with coordinates $(X, Y, Z, 1)^T$ in R_a. From system (7), the ordinate of the projection of this point on the left or right image is $v_l = v_r = v$:

$$v = \frac{[v_0 \sin\theta + \alpha\cos\theta](Y + h) + [v_0\cos\theta - \alpha\sin\theta]Z}{(Y + h)\sin\theta + Z\cos\theta} \tag{8}$$

Moreover, the disparity Δ of the point P is:

$$\Delta = u_l - u_r = \frac{\alpha b}{(Y + h)\sin\theta + Z\cos\theta} \tag{9}$$

From (8) and (9), the plane with the equation $Z = aY + d$ in R_a is projected along the straight line of equation (10) in the "v-disparity" image:

$$\Delta_M = \frac{b}{ah - d}(v - v_0)(a\cos\theta + \sin\theta) + \frac{b}{ah - d}\alpha(a\sin\theta - \cos\theta) \tag{10}$$

N.B: when $a = 0$ in equation (10), we have the equation for the projection of the vertical plane with the equation $Z = d$:

$$\Delta_M = \frac{b}{d}(v_0 - v)\sin\theta + \frac{b}{d}\alpha\cos\theta \tag{11}$$

When $a \to \infty$, we have the equation of the projection of the horizontal plane with the equation $Y = 0$:

$$\Delta_M = \frac{b}{h}(v - v_0)\cos\theta + \frac{b}{h}\alpha\sin\theta \tag{12}$$

3.6 "V-disparity" image construction

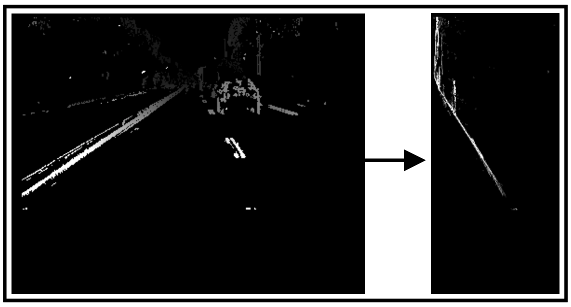

Figure 3: **Construction of the grey level "v-disparity" image from the disparity map. All the pixels from the disparity map are accumulated along scanning lines.**

We suppose that a disparity map I_Δ has been computed from the stereo image pair. For example, this map is computed with respect to the epipolar geometry; the primitives used are horizontal local maxima of the gradient; matching is local and based on normalized correlation (in order to be more robust to global illumination changes) around the local maxima.

Let H be the function of the image variable I_Δ such that $H(I_\Delta) = I_{v\Delta}$. We call $I_{v\Delta}$ as the "v-disparity" image. H accumulates the points with the same disparity that occur on a given image line i. For the image line i, the abscissa u_M of a point M in $I_{v\Delta}$ corresponds to the disparity Δ_M and its grey level i_M to the number of points with the same disparity Δ_M on the line i : $i_M = \sum_{P \in I_\Delta} \delta_{v_P,i}\delta_{\Delta_P,\Delta_M}$ where $\delta_{i,j}$ denotes the Kronecker delta (see Fig. 3).

Once I_Δ has been computed, $I_{v\Delta}$ is built by accumulating the pixels of same disparity in I_Δ along the $\vec{v}$ axis.

3.7 Robust determination of the plane of the road

With the mean values used for the parameters of the stereo sensor, the plane of the road is projected in $I_{v\Delta}$ as a straight line with mean slope 0.70. The longitudinal profile of the road is therefore a straight line in $I_{v\Delta}$. Robust detection of this straight line can be achieved by applying a robust 2D processing to $I_{v\Delta}$. In our application we use a Hough transform, the bounds of Hough space depending on the extreme values of h and θ that are tolerated.

When the road is not planar, the longitudinal profile of the road can also be estimated. The method used is described in details in [1].

3.8 Robust determination of the obstacle location

With the mean values used for the parameters of the stereo sensor, the plane of an obstacle is projected in $I_{v\Delta}$ as a straight line nearly vertical. Thus we just have to extract vertical straight lines in $I_{v\Delta}$ in order to extract obstacles. In this purpose, we build an histogram that accumulates all the grey values of the pixels for each column of the $I_{v\Delta}$ image and then we search for maxima in this histogram.

It is then possible to compute the ordinate of the contact point between the obstacle and the road surface. The distance D between the vehicle and the obstacle then given by:

$$D = \frac{b(\alpha\cos\theta - (v_r - v_0)\sin\theta)}{\Delta} \tag{13}$$

where v_r is the ordinate of the road-obstacle contact point in the image.

3.9 Theoretical good properties of the algorithm

It should be noticed that the algorithm is able to detect any kind of obstacles since the detection is generic. Furthermore, all the information in the disparity map I_Δ is exploited and the accumulation performed increases the density of the alignments in $I_{v\Delta}$. Any matching errors that occur when I_Δ is computed cause few problems as the probability that the points involved will generate coincidental alignments in $I_{v\Delta}$ is low. As a matter of fact, the algorithm is able to perform accurate detection even in the event of a lot of noise or matching errors, and when there is only a few correct matches or a few amount of correct data in the images: in particular in night condition when the majority of the pixels are very dark. Finaly, the algorithm works whatever the robust process used for computing the disparity map or for processing the "v-disparity" image.

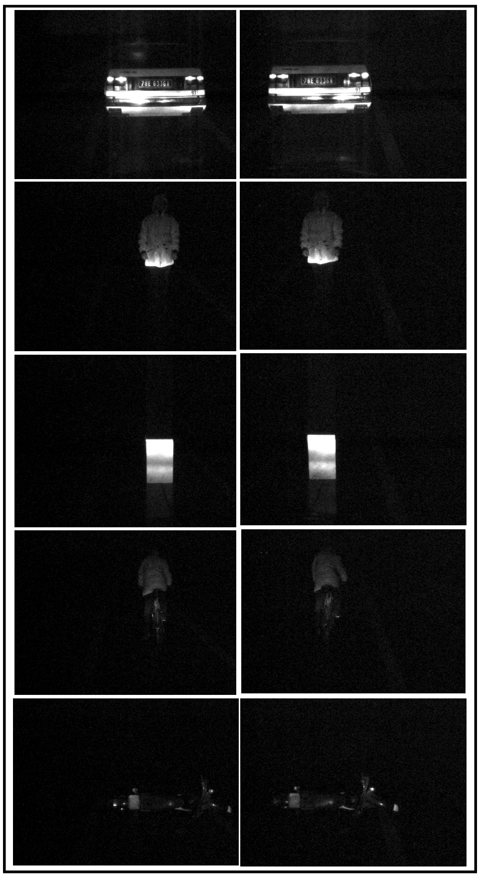

Figure 4: **The five obstacles used for evaluation of distance computation accuracy, located at 5 meters from the vehicle (see text).**

4 DETECTION OF VARIOUS TYPES OF OBSTACLES IN NIGHT CONDITION

Five types of obstacles are used for the evaluation of the distance computation accuracy and for the confidence value computation in night condition: a 1.70 m high pedestrian, a 1.70 m high cyclist, a 1.50 m high vehicle, a fallen motorbike, and a $0.70 \times 0.40 \times 0.40$ m box. Every obstacle is positioned along the axis of the vehicle, at the following distances : 3 m, 5 m, 10 m, 15 m, 20 m, 25 m, 30 m, 35 m, 40 m. Fig. 4 shows a stereoscopic pair of each obstacle in night condition at 5 meters from the stereoscopic sensor. The images are very dark because the exposure time is low.

For each obstacle at every distance, a disparity map is computed (each pixel with horizontal gradient above 3 is taken into account). Then, the "v-disparity" image is computed, the obstacle is detected, the road plane is evaluated, the road-obstacle contact point is computed and its distance from the stereo sensor is computed from (13). It should be noted that the markings of the road and the structures located on the road surface are hardly visible because the images are dark. As a matter of fact, in some cases it is difficult to evaluate the road plane geometry. If it is the case, the straight line corresponding to the road is approximated thanks to the knowledge of the geometric features of the stereo sensor (cf. Section 2), using the equation (12).

Fig. 5 and 6 show the distance computation results. Since the disparity precision is currently one pixel, the two theoretical curves that define the theoretical range (upper and lower) of the distance evaluation are drawn on the figure. Concerning this theoretical range, the maximum difference between the reference distance and the computed one is 0.7% at 3 m and 14% at 40 m. It should be noted that the actual computed distance values are mostly in the theoretical range. The out-of-range values concern distances less than 20 m, where the plane characterisation of the obstacles is not precise enough; it is also the case of the cyclist: indeed it is positionned so that the reference distance measures the wheel position whereas the algorithm measured the back of the cyclist position, which is more visible in the images. However, the error between the reference distance and the computed distance does not exceed 7%. Other stereo systems has been experimentaly evaluted in [3], [4] and [5]. Qualitative evaluation in night condition has been carried out for this last algorithm.

The confidence value is computed by adding all the "v-disparity" grey values for pixels belonging to the same obstacle. In order to avoid any false detection, an obstacle is considered to be detected only if the confidence value is above a threshold set to 20 in our system.

Fig. 7 and 8 show the confidence value results for all the obstacles. This confidence value increases the higher is the obstacle, when the obstacles distance decreases, and when the number of non-vertical edges of the obstacles increases. It should be noticed that the confidence value of the box (respectively the motorbike and the cyclist) is under 20 when the distance is above 20 m (respectively 30 m). This is mainly because the height of the box and the motorbike is low and because they are out of the ligthning area of the car when their distance increase. Concerning the cyclist, it is because the number of non-vertical edges in the images is lower than the one of the other obstacles.

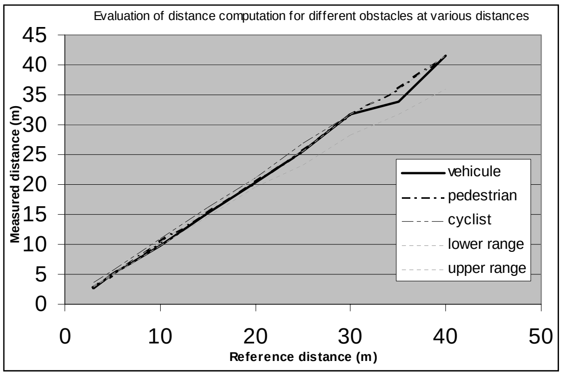

Figure 5: **Distance computation accuracy.**

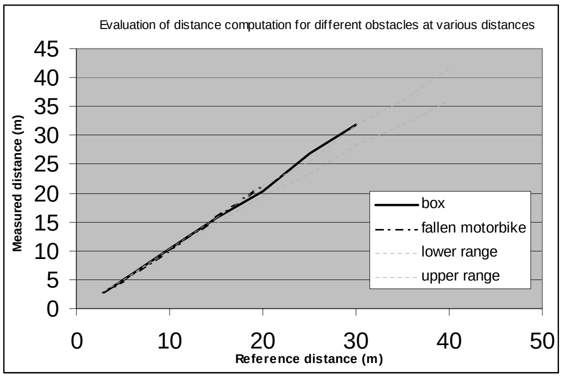

Figure 6: **Distance computation accuracy.**

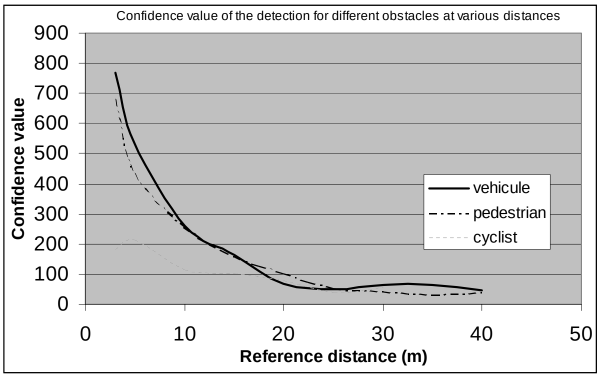

Figure 7: **Confidence value.**

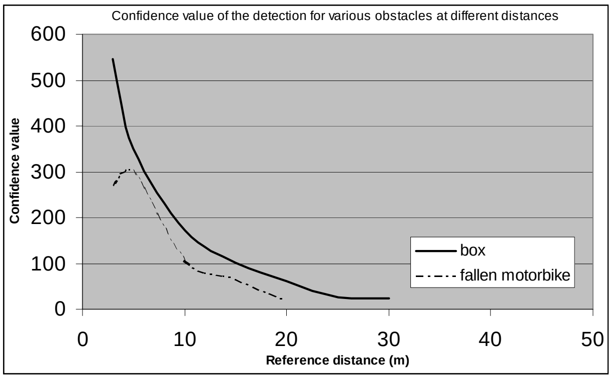

Figure 8: **Confidence value.**

5 DETECTION UNDER ADVERSE CONDITIONS

The system has been tested intensively under adverse meterologic conditions (rain in night condition). Experiments show that the system is merely affected by such conditions. In fact, the results are exactly the same at night whatever is the weather (rainy or not rainy). This is because the rain does not have any effect on the images at the windscreen level when there is no light. The only effect is a mirror effect on the road surface. However this mirror effect does not affect in any manner our system since we look for obstacles above the road surface and since the artefacts caused by the mirror effect are localised under the road in the "v-disparity" image (see Fig. 9).

The system has also been tested intensively under various and adverse meterologic conditions in day condition (glowing effects due to sun, rain) [2]. Experiments show that the system is merely affected by such conditions until the obstacle is visible.

Moreover, experiments of gaussian noise addition and good matches removal (replaced by randomized false matches) in the disparity map show that the "v-disparity" algorithm goes on working efficiently even when there is 97% noise addition or 60% good matches removal (see Fig. 10. top: the reference disparity map and corresponding "v-disparity" image - middle: 97% gaussian noise addition - bottom: 60% good matches removal) for a vehicle located at 20 m, in night and rainy condition. In both cases, relevant information can still be extracted and the obstacle detection process is efficient. These experiments do not represent perfectly the real noise that can affect the disparity map (which is more likely to be correlated noise)

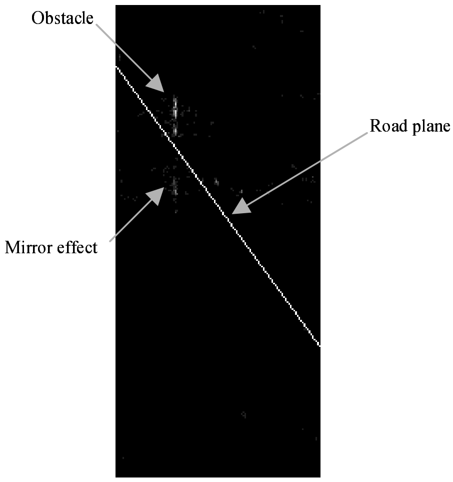

Figure 9: **The ”v-disparity” image for a vehicle located at 20 m in night and rainy condition. There is a mirror effect under the straigth line of the road which does not affect our algorithm since obstacles are researched above the road surface.**

but gives an idea about the robustness of the algorithm. These experiments confirm the ones carried out in day condition [2].

For all the experiments, the tests and the evaluations carried out, computation time does not exceed 40 ms using a $PIII$ 733 GHz PC computer. It should be noted that the computation time is even shorter than the one observed in day condition [2]. This is mainly because there is fewer information in the images in night condition (the images are very dark and the environement is hardly visible). As a matter of fact, there are less pixels to match and the matching process is faster.

6 FUTURE WORKS AND CONCLUSION

This paper has presented an experimental evaluation of the ”v-disparity” algorithm for obstacle detection in night condition. The experimental results confirm the theoretical good properties of the ”v-disparity” algorithm [1]. Distance computation accuracy and confidence have been evaluated. Robustness to noise and adverse conditions has been stressed. Thus, the presented experiments show that stereovision can be used as an efficient perception process for obstacles detection, in night condition

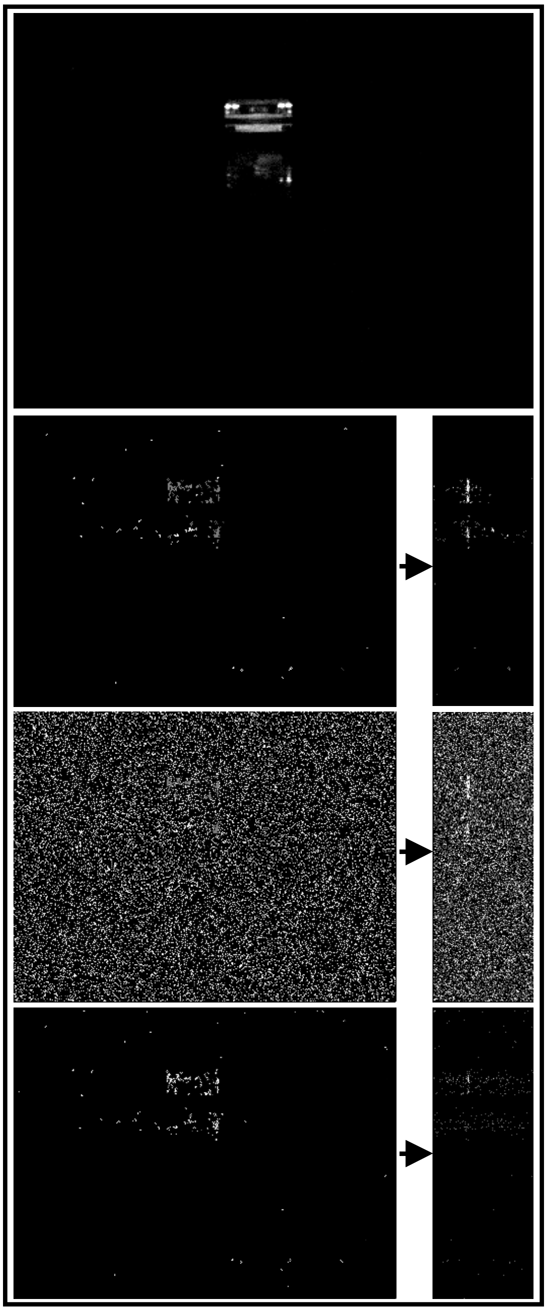

Figure 10: **Top: the original right image of a vehicle at 20 m. Bottom: robustness towards gaussian noise and false matches in the disparity map (see text). Relevant information can still be extracted.**

using CCD cameras, as long as the obstacles are in the lighting area of the car. However the obtained images are very dark. Additionnal tests using an higher exposure time should be carried out. It would be also interesting to evaluate the algorithm performance using CMOS cameras whose dynamic is far higher. The stereo system will be tested as the perception process of an emergency braking and collision avoidance system, in both day and night condition. Future work will also be concerned with the use of the algorithm in a multi-sensors perception application featuring stereovision, laser-scanner and radar data fusion.

ACKNOWLEDGMENTS

The authors would like to acknowledge Pr. Jean Devars's contribution (University of Paris VI). This work is partly founded by the french ARCOS and MICADO projects.

References

[1] R. Labayrade, D. Aubert, J. P. Tarel, ”Real Time Obstacle Detection on Non Flat Road Geometry through V-Disparity Representation”, IEEE Intelligent Vehicules Symposium, Versailles, June 2002.

[2] R. Labayrade, D. Aubert, ”Robust and Fast Stereovision Based Road Obstacles Detection for Driving Safety Assistance”, Machine Vision and Application 2002 (MVA 2002), Nara, Japan, 11-13 December 2002, pp 624 - 627.

[3] M. Bertozzi, A. Broggi - ”GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection”, IEEE Transaction on image processing, Vol. 7, N1, January 1998.

[4] T.A. Williamson - ”A high-performance stereo vision system for obstacle detection ”, Phd, Carnegie Mellon University, September 1998.

[5] G. Toulminet, A. Bensrhair, S. Mousset, A. Broggi, P. Mich, ”Systeme de stereovision pour la detection d'obstacles et de vehicule temps reel”. In Procs. 18th Symposium GRETSI'01 on Signal and Image Processing, Toulouse, France, September 2001.