HAL
open science

Obstacles Detection on a Road by Dense Stereovision with 1D Correlation Windows and Fuzzy Filtering

Sébastien Lefebvre, Sébastien Ambellouis, François Cabestaing

To cite this version:

Sébastien Lefebvre, Sébastien Ambellouis, François Cabestaing. Obstacles Detection on a Road by Dense Stereovision with 1D Correlation Windows and Fuzzy Filtering. Intelligent Transportation Systems Conference, ITSC '06, Sep 2006, Toronto, Canada. pp.739-744, 10.1109/ITSC.2006.1706830 . hal-00521114

HAL Id: hal-00521114

https://hal.science/hal-00521114

Submitted on 25 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Obstacles detection on a road by dense stereovision with 1D correlation windows and fuzzy filtering

Sébastien Lefebvre, Sébastien Ambellouis
INRETS-LEOST,
20, rue Elisée Reclus
59666 Villeneuve d'Ascq, FRANCE
{sebastien.lefebvre,
sebastien.ambellouis}@inrets.fr

François Cabestaing
LAGIS - UMR CNRS 8146
Bâtiment P2 - USTL, Cité Scientifique
59655 Villeneuve d'Ascq, FRANCE
fcab@ieee.org

***Abstract*— In this paper, we propose an original approach to obstacles detection based on stereovision with mono-dimensional correlation windows. The result of the algorithm is a dense disparity map associated with a confidence map. For each pixel, correlation indices are computed for several widths of windows and several positions of the window centre. Three criteria, extracted from each correlation curve, are combined by a fuzzy filter to define a confidence measure. Our 1D method is compared to a classical 2D method and shows better results in term of errors and density rate. In the context of obstacle detection, we show that a basic segmentation of our disparity map yields a better detection and marking of the obstacles. The method is validated on synthetic image sequences and our results are compared with those obtained using a classical 2D method.**

I. INTRODUCTION

Obstacle detection is an essential problem in intelligent transportation and a lot of research projects have been focused on this topic for several decades. Although obstacle detection seems to be easy with a RADAR based-system, it is not obvious to distinguish between vehicles and bridges or other elements of the road infrastructure. In many studies, researchers have proposed to combine different types of sensors, such as RADAR, lidar, and video-based sensors, to yield complementary and redundant information [1].

For vision-based obstacle detection on a road, the system can use either a single camera [2] [3] or several cameras for stereovision [1] [4]. A stereovision-based system can theoretically yield an accurate and detailed 3D representation of the environment around a vehicle, with a passive sensor at relatively low cost. Previous applications for intelligent vehicles have mostly used sparse, feature-based approaches to stereovision. In Nedevschi *et al.* [5], only a subset of image pixels (e.g. objects edges) are matched, in order to comply with real time processing requirements. However, sparse disparity maps often lack information which would allow object segmentation, classification or other robust high-level treatments. Dense stereovision can improve the processing chain significantly, but is computationaly expensive.

In this paper, we propose an original dense stereo matching algorithm. We focus on developing a low-level algorithm mainly based on a fuzzy combination of the correlation measures to provide confidence measures. The density of the disparity map depends on a threshold applied to confidence measures. Our goal is to provide a dense or semi-dense disparity map with high confidence, thus very few errors.

In the following section, we describe some window-based stereovision techniques ; so, we present specific scene properties related to the transportation environment. Section three describes our algorithm defining (1) the matching method used to determine the correlation indices, (2) the fuzzy logic filters and (3) the decision-making process. In the fourth section, the method is validated in obstacle detection context. This section is devoted to the obstacle extraction from the road and presents results on synthetic stereo image pairs. Then, we conclude this paper and present some future works.

II. STEREOVISION

A. Window-based stereo matching

Stereovision is a typical problem in computer vision and in 3D reconstruction. Given two images — called left and right images — capturing a scene at the same time from two points of view, stereo techniques aim at defining conjugate pairs of pixels, one in each image, that correspond to the same spatial point in the scene. The difference between positions of conjugate pixels, called the disparity, yields the depth of the point in the 3D scene.

Sparse or dense disparity maps can be computed. Sparse disparity maps are produced by stereovision algorithms matching features like edges or corners. Dense disparity maps are computed by algorithms based on the analysis of the grey levels or colours of pixels in a neighbourhood centred on each pixel of the images. In both cases, because of the epipolar geometry, the conjugate of a pixel in the first image lies on the corresponding epipolar line in the second image. This allows one to restrict the matching process to epipolar lines.

In the following sections, we focus our attention on dense stereo approaches based on the analysis of a pixel neighbourhood. In these methods, the pairs of conjugate pixels are computed by maximising (or minimising) a similarity (dissimilarity) function. The properties of this function depend on the grey level distributions in the two windows centred on pixels of epipolar lines. On the one hand, the selection of window shape and size is crucial for the success of matching. The dense stereo algorithms assume that all pixels within the windows have the same disparity. Therefore, the windows

must not be too large. On the other hand, if the window is too small, the data available to estimate the similarity function are not sufficient. In literature, several window-based approaches have been presented to deal with this trade-off.

Kanade and Okutomi have proposed an adaptive neighbourhood method [6], in which they iteratively modify the neighbourhood size and shape according to the local variation of the intensity and current depth estimates. Pérez *et al.* also use a neighbourhood with variable shape determined by analysing the similarity between pixels [7]. Some authors use shiftable windows methods: Fusiello *et al.* have presented a Symmetric Multi-Window (SMW) algorithm [8]. They compute the SSD (Sum of Squared Differences) on nine rectangular windows in which the current pixel is positioned at different places, and keep the window with the smaller SSD. Hirschmüller combines the correlation computed on the window centred on the considered pixel with the correlations computed on several support windows [9].

In these algorithms, the modification of window configuration is based on one or several parameters defining the confidence granted to the matching. This notion is essential in some applications, especially for multisensor fusion algorithm and several confidence measures have been described in literature [10]. In [9], the author analyses the relative difference between the two lowest minima of the correlation function. If this value is higher than a fixed threshold, the window configuration is validated and the confidence is high. In [11], the matching is marked as good if the global minimum of the correlation function is sharp. If the confidence is low, several authors prefer to mark the pixels as unmatched and provide semi-dense disparity maps.

B. 1D vs 2D correlation window

All the previously cited authors describe methods assuming that the images have been rectified [12]. In this case, the conjugate pairs of pixels are located on a single and same line in the left and right image. Usually, authors compute correlation indices with two-dimensional windows shifted along the raster lines to deal with textureless area and increase the density of the disparity map. Unfortunatly, when neighbour pixels within the window don't have the same disparity, errors appear.

Usual road scenes are characterised by a strong perspective with a vanishing point. Thus, two consecutive raster lines correspond to road sections located at slightly different distances from the camera. Therefore, a correlation window including several raster lines does not satisfy the constraint of constant disparity along the vertical direction.

In our work, we assume that the information located on the epipolar line, e.g. the single raster line, is sufficient to provide a good matching. In the following section, we present an original stereo algorithm using 1D correlation windows and a fuzzy logic filter. Each pixel is labeled with a disparity and the associated confidence value.

III. ALGORITHM

In the following, the overall algorithm can be decomposed into three steps :

1) two rectified images are processed using several 1D windows to yield similarity measures for each pixel ;
2) the similarity measures are processed by fuzzy logic filters which associate a confidence value to each of them ;
3) a decision-making process analyses the whole data and selects the most confident disparity for each pixel.

A. Matching method

In correlation-based stereo matching, many techniques use 2D neighbourhoods to increase the available information, but one can discuss the legitimacy of this strategy. With our algorithm, we will show that in many situations, a 1D window is sufficient to find a correct matching — assuming that the raster lines coincide with epipolar lines.

The aim of an area-based dense stereo matching algorithm is to determine correspondences between all the pixels in the two images and to compute the associated disparities. A 1D window of width $2w+1$ is positioned at the pixel (x, y) in the first image. The correlation index is computed for several values of the shift $s \in [1, s_{max}]$, for widths $w \in [1, w_{max}]$ and for three positions p of the current pixel (left, center and right) in the window.

The correlation index $C(p, w, x, y, s)$ is based on the SAD dissimilarity index (Sum of Absolute Differences):

$$C(p,w,x,y,s) = \sum_{i=-w+p}^{i=w+p} |I_l(x+i,y) - I_r(x+i+s,y)| \quad (1)$$

$$with$$

$$\begin{cases} \text{p = w} & \text{if left position} \\ \text{p = 0} & \text{if center position} \\ \text{p = -w} & \text{if right position} \end{cases},$$

where I_l and I_r denote respectively the grey levels of the left and right images. For a given value of p and w, a correlation curve is the list of values $C(p, w, s)$ with $s \in [1, s_{max}]$:

$$\{C(p,w,s)\}_{s\in[1,s_{max}]} \ . \quad (2)$$

In the following, to simplify the notations, the parameters x and y have been removed from the equations.

We have selected the SAD dissimilarity index, despite its well known drawbacks, for simplicity and as a basis of comparison.

In area-based stereo matching, the main problem resides in the determination of the optimal correlation window. The window may be too small, containing not enough information, or too large, including several areas with different disparities. Our algorithm computes correlation curves for w_{max} window sizes and three positions of the current pixel in each window. In the next section, we show how these $3 \times w_{max}$ curves are combined using a set of fuzzy filters.

B. Fuzzy logic filter

For each pixel, the $3 \times w_{max}$ curves are processed by $3 \times w_{max}$ identical fuzzy filters which yield the fuzzy confidences $Conf_{FL}(p, w, s)$ for every position p of the current pixel in the window of size w and for every shift s. In one correlation curve, every shift value is associated with a confidence value which evaluates the likelihood that this shift is the disparity for the current pixel.

We extract three characteristics from the correlation curves:

- The curvature metric $Cur(p, w, s)$ [10], which measures the curvature on the correlation curve $C(p, w, s)$ at shift s (excluding border effects) ; a sharp correlation valley and a high value of $Cur(p, w, s)$ characterise an accurate matching and mark the disparity with low ambiguity. $Cur(p, w, s)$ is defined as:

$$\begin{aligned} Cur(p, w, s) = & \quad -2 \cdot C(p, w, s) \\ & +C(p, w, s+1) + C(p, w, s-1) \ . \end{aligned}$$

- The rank $R(p, w, s)$: Correlation values $C(p, w, s)$ are sorted in increasing order, and $R(p, w, s)$ denotes the rank of $C(p, w, s)$ in this list. Since the SAD minimises the similarity function, there is a strong likelihood of having mimima of the curve in the first ranks.
- The number $N(p, w)$ of inflexion points characterizing minimum peaks (convex curves) which are defined by:

$$\frac{\partial^2 C(p, w, s)}{\partial s^2} = 0 \ .$$

These three characteristics are described by fuzzy sets with 3 states (bad, medium, good), defined by classical membership functions (triangular and trapezoidal). For example:

- when $Cur(p, w, s)$ is high the state of the membership function is *good*
- when $N(p, w)$ and $R(p, w, s)$ are high the state of their membership functions is *bad*

In the same way, the output $Conf_{FL}(p, w, s)$ of the fuzzy filter is described by a variable with five states: null, bad, medium, good, excellent. We use a regular IF-THEN-ELSE structure as the inference mechanism. The 27 inference rules have been defined by analysing .the behaviour of the three characteristics, Cur, R, and N on typical neighborhoods. Some rules, like the following one, are obvious:

IF **Cur(p,w,s)** is '*good*',
AND **R(p,w,s)** is '*good*',
AND **N(p,w)** is '*good*'
THEN **Conf$_{\mathbf{FL}}$(p,w,s)** is '*excellent*'.

However, for many rules, it is difficult to define precisely the relationship between the fuzzy variables. For example, for two values s_1 and s_2, if the corresponding curve shows that $(Cur(p, w, s_1) = good\ AND\ R(p, w, s_1) = bad)$ and $(Cur(p, w, s_2) = medium\ AND\ R(p, w, s_2) = good)$, it is not obvious to define which one will result in the higher confidence (excellent). In the present case, the associated confidences are characterised by the medium state : the decision step will be determining.

Each rule computes an elementary confidence for a given value of p, w and s. In the deffuzification step, the 27 confidence values are combined using the center of gravity method, to yield the global confidence value $Conf_{FL}(p, w, s)$.

A confidence value $Conf_{LF}(p, w, s)$ is associated with every shift of every correlation window, then a decision-making process determines the disparity corresponding to the best matching for each pixel (x, y).

C. Decision-making process

Firstly, we select the shift $s^*(p, w)$ corresponding to the maximum confidence value for each configuration of the window (width and position):

$$s^*(p, w) = \arg[\max_s \{Conf_{LF}(p, w, s)\}] \ . \tag{3}$$

Then, we compute the number $N_{occ}(s)$ of occurrences of these maximum values as:

$$N_{occ}(s) = Count_{p,w}[s = s^*(p, w)] \ . \tag{4}$$

Then, we determine the disparity d as the shift value which has the maximum number of occurrences in all the configurations:

$$d = \arg[\max_s (N_{occ}(s))] \ . \tag{5}$$

Let $\{Conf_{FL}(p, w, s^*(p, w))\}_{s^*(p,w)=d}$ be the set of values $Conf_{FL}(p, w, s^*(p, w))$ for which $s^*(p, w) = d$. To the disparity d, we associate a confidence which is the minimum value of this set for all values of p and w:

$$Conf^*_{FL} = \min_{p,w} \{Conf_{FL}(p, w, d)\}_{s^*(p,w)=d} \ . \tag{6}$$

For some neighbourhoods, this confidence value is not very high, but all window configurations yield the same disparity. In this case, we consider that the confidence associated to d must be increased. Indeed, to take into account the number $N_{occ}(d)$ of occurrences of the disparity d among all window configurations, we define a second confidence value $Conf_{occ}$ as following:

$$Conf_{occ} = \frac{N_{occ}(d)}{3 \times w_{max}} \ . \tag{7}$$

The final degree of confidence $Conf$ associated with the selected disparity d is computed as:

$$Conf = \frac{\alpha \cdot Conf^*_{FL} + \beta \cdot Conf_{occ}}{\alpha + \beta} \ , \tag{8}$$

where α and β are a priori defined weighting coefficients.

The decision-making process yields a disparity and an associated confidence value for every pixel of the reference image. Finally, a semi-dense disparity map can be composed of the pixels whose confidence is higher than a threshold. The lower the threshold, the denser the disparity map.

Although our method reduces the overall rate of incorrect matchings, disparities with low confidences remain in some

areas. In [10], the authors apply the left-right consistency (LRC) constraint to further remove incorrect matchings. Initially, the proposed method provides, for each pixel of the left image, the best match into the right image. By reversing the role of both images, for each pixel of the right image we find the best match in the left image. Finally, we keep only the matches that turn out to be coherent when matching left-to-right (direct matching phase) and right-to-left (reverse matching phase). Formally, if the pixel with coordinate x_R in the right image matches the pixel with coordinate $x'_L = x_R + d^R_{x_R}$ in the left image ($d^R_{x_R}$ is said to be the right-based horizontal disparity), then LRC is defined as:

$$Error = x_R - (x'_L + d^L_{x'_L}),$$

where $d^L_{x'_L}$ is the left-based horizontal disparity. When the error is greater than zero, left and right-based matches disagree and the computed disparity is removed and set to zero on the disparity and confidence maps.

IV. RESULTS

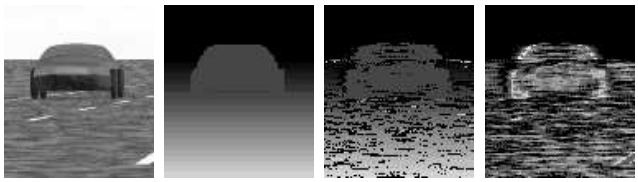
In the following section, different results are presented and compared : the results based on our 1D method and those based on the classical 2D Hirschmüller method [9]. Firstly, we will compare the computed disparity maps. Then, we apply a basic algorithm for road and obstacle detection on both 2D and 1D disparity map and we compare the quality of the detections.

A. Comparison of disparity maps

We have evaluated our stereovision technique on a synthetic image sequence of a road scene generated with POV-Ray^{TM} rendering software. We will present the obtained results on the pair of images 10 of the sequence.

Our 1D method yields a disparity map associated with a confidence map. The disparity map is computed using 1D windows with sizes ranging from 3×1 to 11×1 (i.e. $w \in \{1..5\}$ and $w_{max} = 5$). Small windows are retained in order to test theirs capabilities to make good matching. For equation (8), several values of α and β have been tested and the better results have been obtained with α and β equal to one. The figures 1(a), 1(b), 1(c) and 1(d) present respectively the same piece extracted from the left image, from the ground truth, from the disparity map and from the associated confidence map. In the disparity map, a white pixel represents a large disparity (objects close to the camera) and a black one is either a small disparity (objects far from the camera) or a cancelled disparity. In the confidence map, a white pixel marks a disparity with high confidence, a black pixel a disparity with low confidence. As the figure 1(c) shows, our algorithm yields good results with few errors and having high density.

We compare our method with the Hirschmüller 2D window-based method [9]. Hirschmüller combines the correlation computed on the window centered on the considered pixel with the correlations computed on several support windows [9]. In this paper, we select the configuration with one window in the middle surrounded by four partly overlapping windows. Firstly, the correlation values are computed with this configuration. Afterward, the left-right consistency check invalidates places of uncertainty. Disparity maps have been computed for windows sizes ranging from 3×3 to 21×21.



(a) (b) (c) (d)

Fig. 1. (a) Left image, (b) Ground truth, (c) Disparity map, (d) Confidence map

The quantitative comparison is based on the density and the error rate computed for both disparity maps. Firstly, error rate is computed by the absolute pixel to pixel difference $|\Delta\varepsilon|$ between the resulting dense disparity map (obtained with our 1D method and the Hirschmüller method) and the ground truth. By counting the number of validated pixels, we compute the error rates ε_1 and ε_0 for respectively $\Delta\varepsilon$ greater than one and zero. Density is obtained by counting the number of validated disparities for both disparity maps for all marked disparities of the ground truth.

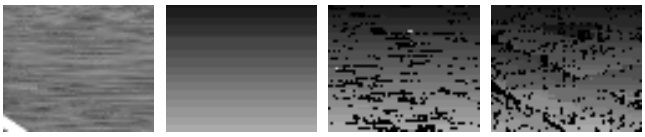
Results are presented in table I. In the 2D window case, we keep the better results obtained for a 2D correlation window with size 9×9. One can notice that with a higher density, our 1D method yields better results whatever the error threshold choosen (1 or 0).

TABLE I
COMPARISON OF ERROR RATES

Method	ε_0 $\lvert\Delta\varepsilon\rvert > 0$	ε_1 $\lvert\Delta\varepsilon\rvert > 1$	Density
Hirschmüller (window 9×9)	18.48 %	7.30 %	59.55 %
1D Method	9.88 %	2.53 %	70.47 %

In the same manner, a piece of the left image characterising exclusively the road is computed with the two algorithms. The figures 2(c) and 2(d) present respectively the disparity maps computed with our 1D method and the Hirschmüller method. In comparison with the ground truth (Fig. 2(b)), one can notice that the generated disparity ramp, obtained with our method, seems closer to the ground truth. Also, the error percentages ε_1 and ε_0 have been computed and are presented in table II.

One can notice that for the same density, our 1D method yields better results whatever the error threshold choosen. The error rates associated with the 2D method are very bad. The method can not yield precise disparity values in this situation. This is due to the fact that two consecutive



(a) (b) (c) (d)

Fig. 2. (a) Left image, (b) Ground truth, (c) Disparity map computed with the 1D method, (d) Disparity map computed with the Hirschmüller method

TABLE II

COMPARISON OF ERROR RATES ON THE ROAD SECTION

Method	ε_0 $\lvert\Delta\varepsilon\rvert > 0$	ε_1 $\lvert\Delta\varepsilon\rvert > 1$	Density
Hirschmüller (window 9×9)	48.83 %	32.85 %	82.32 %
1D Method	12.41 %	1.25 %	82.84 %

lines correspond to road sections located at slightly different distances from the camera. Therefore, the correlation window including several lines does not satisfy the constraint of constant disparity along the vertical direction and provides a lot of matching errors.

B. Road and obstacle detection

In this section, we present a ground plane estimation and the obstacle detection in a typical application of stereovision in intelligent transportation system. Road and obstacle detection by analysing a disparity map is a two steps method. The first step defines which pixels correspond to the vehicles and which ones correspond to the road. The second has to group the pixels in order to detect and separate each obstacle in front of the stereo sensor.

Different methods dealing with estimation of the ground plane are described in the literature. Some of them, as the approach of Labayrade *et al.* [4], are based on the analysis of the histogram of disparity values. This is a robust method for detecting the obstacles which is based on the assumption that, for each scanline where the road is visible, the dominant disparity value is that of road surface pixels. A "V-disparity" image is computed for each scanline. A "V-disparity" is the histogram of disparity values of the associated scanline in the original disparity image. The ground plane is estimated by extracting from this "V-disparity" image the dominant line features with the Hough transform.

Other methods [13] consider that a road and an obstacle can be approximated respectively by an horizontal and a vertical plane in the real scene. As it is described in [13], road disparity $d_{road}(x, y)$ is a linear function in terms of disparity map pixels coordinates:

$$d_{road}(x, y) = ax + by + c \ , \tag{9}$$

where (a, b, c) are the parameters of the ground plane and (x, y) are pixel coordinates.

Under this assumption, a pixel corresponds to an obstacle if its disparity value satisfies the following equation:

$$d(x, y) > ax + by + c \ . \tag{10}$$

In the literature, authors generally describe a two steps method. Firstly, they compute the ground plane parameters by fitting a plane in the (x, y, d) space using a least squares method. Secondly, they verify that each pixel of the disparity map satisfies inequality (10).

Under the same assumption, we modify this classical method by analysing separately the profile of every column in the disparity map. For each column x of the disparity map, the disparity profile of road pixels is a straight line segment with a large slope. On the contrary, for the pixels corresponding to an obstacle, the slope of the straight line is small. Thus, boundaries of an obstacle are marked at the intersections of two line segments in the profile, one with a small slope and one with a large slope. Bounding boxes are generated using the edges detected on neighbour columns with the constraint of constant disparity to define each bounding box.

One can apply this obstacle detection method on the disparity maps computed with our 1D method and with the Hirschmüller one : the associated results are shown on the figures 3(d) and 3(f) for which one can visualise the bounding boxes superimposed on the left image. Fig. 3(b) represents the bounding boxes generated on the ground truth for basis of comparison. Our algorithm yields very good results : the obstacles are clearly detected even when they are far from the stereo sensor. In fig. 3(f), one can notice that the Hirschmüller method does not provide efficient bounding boxes. In this way, they appear around pixels characterising a road section and bring false detections.

Also, with the classical 2D method, many errors are located at discontinuities in the disparity map, i.e. at obstacle boundaries. Therefore, the bounding boxes are not precise. With our method, errors do not appear specifically near discontinuities, and the bounding boxes are closer to obstacles boundaries. Thus, the models used in a subsequent classification or tracking stage are more accurate. Finally, we can note that obstacles are well separated thanks to the constraint of constant disparity used to define each bounding box.

V. CONCLUSIONS AND FUTURE WORKS

This paper introduces an original approach to correlation-based stereo matching. For every pixel, similarity measures are evaluated with several 1D correlation windows of different widths, in which the current pixel is not necessary placed at the center. The correlation indices are further processed by a set of fuzzy filters to assign a confidence to every shift value. A decision-making process analyses the confidence values for all window widths and positions to determine a disparity for each pixel and an associated confidence value. The method has shown the ability to provide good matching in most cases particularly for the pixels characterising the road. The algorithm takes the advantage of the 1D property

and yields much better results than the classical 2D method. In a road and obstacle detection context, our method yields good results because the disparity map is quite dense with very few errors. The Hirschmüller method brings a lot of false detection due to the average effect of the 2D correlation windows.

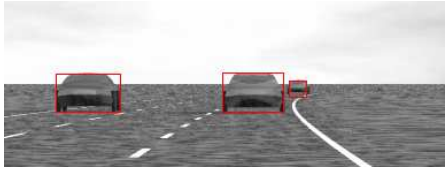
The proposed method includes only low level algorithms, i.e. that can be applied independently and with the same parameters on every image neighbourhood. Its advantages are that (1) it is fundamentally parallel due to the correlation window shape and to the structure of the fuzzy logic filters and that (2) no parameter tuning is required. Of course, with a standard sequential implementation, computation time is high and incompatible with the real-time constraint. However, the method can be implemented in real time on dedicated hardware.

At the moment, experiments are carried out on real world scenes, for which the rectification stage is crucial since the method only uses single raster lines for pixel matching. In future works, for real-time applications, we will implement the algorithm on dedicated architecture hardware based on a FPGA from Altera, called the STREAM processor, that has been developed by LAGIS and LEOST. We will improve obstacle detection and tracking by using a lane-markings extraction process.

VI. ACKNOWLEDGMENTS

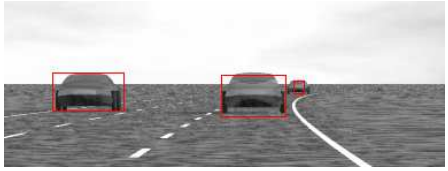
This work is part of the french research project RaViOLi, supported by EU, french government and the Nord-Pas-de-Calais regional council under contract number OBJ2-2005/3-4.1-253-7820.

(a) Ground truth



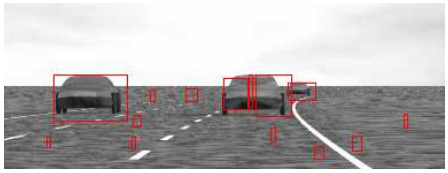
(b) Left image with bounding boxes based on (a)

(c) Disparity map with 1D method



(d) Left image with bounding boxes based on (c)

(e) Disparity map with Hirschmüller method (size 9×9)



(f) Left image with bounding boxes based on (e)

Fig. 3. Results on images pair 10.

REFERENCES

[1] C. Stiller, J. Hipp, C. Rössig, and A. Ewald, "Multisensor obstacle detection and tracking," *Image and vision computing*, vol. 18, no. 5, pp. 389–396, Apr. 2000.

[2] B. Steux, C. Laurgeau, L. Salesse, and D. Wautier, "Fade : A vehicle detection and tracking system feturing monocular color vision and radar data fusion," in *Procs. IEEE Intelligent Vehicles Symposium 2002*, Versailles, France, June 2002.

[3] A. Broggi, P. Cerri, and P. C. Antonello, "Multi-resolution vehicle detection using artificial vision," in *Procs. IEEE Intelligent Vehicles Symposium 2004*, Parma, Italy, June 2004, pp. 310–314.

[4] R. Labayrade, D. Aubert, and J. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through V-disparity representation," in *Procs. IEEE Intelligent Vehicles Symposium 2002*, Versailles, France, June 2002, pp. 646–651.

[5] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, R. Schmidt, and T. Graf, "High accuracy stereo vision system for far distance obstacle detection," in *Procs. IEEE Intelligent Vehicles Symposium 2004*, 2004, pp. 292–297.

[6] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: theory and experiment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 9, pp. 920–932, Sept. 1994.

[7] M. Pérez, F. Cabestaing, O. Colot, and P. Bonnet, "A similarity-based adaptive neighborhood method for correlation-based stereo matching," in *IEEE International Conference on Image Processing*, Singapore, 2004, pp. 1341–1344.

[8] A. Fusiello, V. Roberto, and E. Trucco, "Symmetric stereo with multiple windowing," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, no. 8, pp. 1053–1066, 2000.

[9] H. Hirschmüller, "Improvements in real-time correlation-based stereo vision," in *IEEE Workshop on Stereo and Multi-Baseline Vision*, Kauai, HI, USA, Dec. 2001.

[10] G. Egnal, M. Mintz, and P. Wildes, "A stereo confidence metric using single view imagery with comparison to five alternative approaches," *Image and vision computing*, vol. 22, no. 12, pp. 943–957, Oct. 2004.

[11] K. Mühlmann, D. Maier, J. Hesser, and R. Männer, "Calculating dense disparity maps from color stereo images, an efficient implementation," in *IEEE Workshop on Stereo and Multi-Baseline Vision*, Kauai, HI, USA, June 2001, pp. 30–36.

[12] A. Fusiello, E. Trucco, and A. Verri, "Rectification with unconstrained stereo geometry," in *Proceedings of the British Machine Vision Conference*, Essex, U.K., Sept. 1997, pp. 400–409.

[13] S. Se and M. Brady, "Ground plane estimation, error analysis and applications," *Robotics and Autonomous Systems*, no. 39, pp. 59–71, 2002.