A colour correlation-based stereo matching using 1D windows

Sébastien Lefebvre, Sébastien Ambellouis, François Cabestaing

HAL Id: hal-00521116
https://hal.science/hal-00521116

Submitted on 25 Sep 2010

A colour correlation-based stereo matching based on 1D windows

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 colour correlation-based stereo matching with mono-dimensional windows. The result of the algorithm is a quasi-dense disparity map associated with its confidence map. For each pixel, correlation indices are computed for several widths of windows and several positions of the current pixel. Three criteria, extracted from each correlation curve, are combined by a fuzzy filter to define a confidence measure. A basic decision rule computes the disparity value and its associated confidence for the most of the image pixels. A first study shows results obtained on grey level images with our 1D method and a classical 2D method. Here, the 1D approach presents better results. Moreover, our method is applied to the RGB colour space. A disparity map is computed with each of the three colour components. A fusion step allows to compute the disparity values based on these three disparity maps. The method is validated on the Tsukuba image pair. In the first time, we show that our method presents lower error rates with the RGB colour space than with the grey level image and this, for identical density rates. In a second time, our results (with the colour way) are compared in terms of errors and density rate with those obtained using similar colour 2D methods (presented on the Middlebury website). Our algorithm is ranked in the first places for each area of the image.

I. INTRODUCTION

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 or colour 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 [1], 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 [2]. Some authors use shiftable windows methods: Fusiello *et al.* have presented a Symmetric Multi-Window (SMW) algorithm [3]. 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 [4].

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 [5]. In [4], 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 [6], 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 [7]. 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. Unfortunately, when neighbour pixels within the window do not have the same disparity, errors appear.

In images of real scenes, this fronto-parallel plane assumption — i.e. assumption that all the pixels have the same disparity — is regularly violated because of the perspective effect or because complex objects cannot be accurately described with planes. In this case, disparity errors and a blurring effect across depth discontinuities appear in the resulting map. Moreover, the larger the size of the 2D window, the higher the error rate.

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 sections, we will 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. A basic decision step determines the disparity value and its confidence for the most of image pixels. Indeed, in the case of occlusions or in untextured areas, no information is locally available and most area-based methods compute wrong disparities. To minimize the effect of this drawback, we propose a quasi-dense method.

C. Colour correlation-based matching

In computer vision application, the choice of colour space is important. However, a lot of colour spaces have been proposed; In [8], Vandenbroucke *et al.* have classified the most of the used colour spaces. Several authors have used colour information for correlation-based stereo matching algorithm. Indeed, as presented in [9], the use of colour images can improve the accuracy of stereo matching. The authors have tested their stereo matching method with nine different colour spaces: The results show that colour always improves matching however the best colour space is not easy to underline.

In this paper, we aim at taking into account the colour information with our 1D grey level matching approach. We compute a disparity map with each colour component and combine the results in order to have a denser and more precise disparity map.

The paper is organised as follows. In Section 2, our 1D windows-based approach is described. Section 3 presents the first results obtained on grey level image; A comparison is done with results computed with the classical 2D method of Hirschmüller using the Middlebury evaluation framework. In Section 4, we introduce the use of colour information by detailing the combination method. The colour way is compared to the grey level one. So, our disparity map is compared to maps computed by four other similar colour methods related in literature. The last Section is dedicated to conclusions and outlines future studies.

II. THE 1D APPROACH

The proposed method relies on a particular combination of correlation scores computed on several 1D windows. It provides a disparity map associated to a confidence map. The confidence map indicates the level of certainty of each matching, which can be low for example in untextured regions. The overall method is composed of the four steps presented in figure 1. The following sections describe each step, except the left and right consistency (LRC). LRC is a very classical method used to remove incorrect matchings that has already been widely presented in the literature. The reader can find a detailed explanation of this technique in [5].

In the figure 1, the correlation scores, the confidence values and the final disparity map can be computed using various tools and theories. In this part, we present a case study in which the most basic techniques are used for each stage.
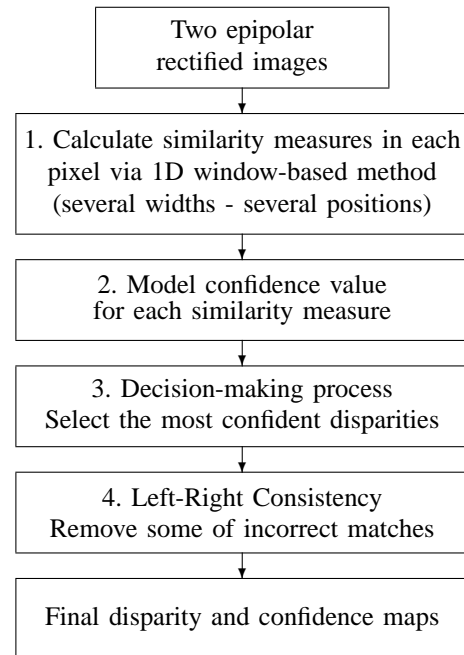

Fig. 1. Algorithm outlines.

A. Similarity measure

Classical area-based dense stereo algorithms assume that all pixels within correlation windows have the same disparity. Determining the optimal correlation window is crucial for the success of matching, but doing this a priori without specific knowledge about the scene is often impossible. Selecting a specific window size and shape corresponds to making a decision at the beginning of the process. We propose to postpone the decision until the last stage of the matching process. Thus, we compute 1D correlation scores for several widths of the window and several positions of the current pixel.

Let $w_n = 2n+1$ be the width of a 1D window Ω_n, where n is an integer in the range $[1, n_{max}]$. The window corresponds

to a neighbourhood of the reference pixel (x, y). Let p_i be the position of the reference pixel in Ω_n, with $p_i \in [-n..n]$. An identical window is placed in the other image at pixel (x, y) and shifted by an integer value s within range $[s_{min}, s_{max}]$.

Figure 2 represents what we call the volume of all correlation values computed for a single reference pixel. Each cube represents the correlation value computed on a 1D window with width w_n ($w_n \in [w_1, w_{n_{max}}]$) shifted by s and for which the position of the current pixel is p_i. Each line of cubes in the volume along the s direction is equivalent to a correlation curve obtained for a given position of the reference pixel in the correlation window and for a given width of this window.
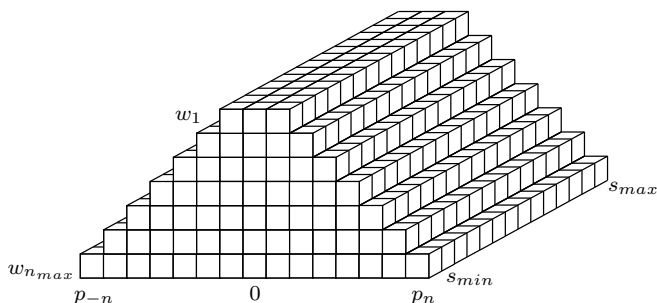

Fig. 2. Correlation volume, for a single reference pixel, defined by the whole set of configurations of the correlation window

In order to show that our 1D approach performs well whatever correlation technique is chosen, we have selected a basic one, i.e. the Sum of Squared Differences (SSD) dissimilarity index. Thus, the correlation index $C(x, y, p, w_n, s)$ is defined as:

$$C(x, y, p, w_n, s) = \sum_{i=-p+n}^{i=-p-n} (I_l(x+i, y) \qquad (1)$$
$$-I_r(x+i-s, y))^2$$

$$\text{with} \quad \begin{cases} w_n = 2n + 1 \\ n \in [0, n_{max}] \\ p \in [-n, n] \\ s \in [s_{min}, s_{max}] \end{cases}, \qquad (2)$$

where I_l and I_r denote respectively the grey level functions of the left and right images.

For a given value of p and w_n, the correlation curve is defined by the list of values $C(p, w_n, s)$ with $s \in [0, s_{max}]$:

$$\{C(p, w_n, s)\}_{s \in [0, s_{max}]} . \qquad (3)$$

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

Correlation curves are computed using 1D windows with sizes ranging from 3×1 to 21×1 (i.e. $n \in [1, 10]$ and $n_{max} = 10$). Small windows are retained in order to test their capabilities to provide a good matching.

In area-based stereo matching, determining the optimal correlation window is the main problem. We propose to analyse the correlation scores for different windows: we compute

correlation curves for n_{max} window sizes and $(2n_{max} + 1)$ positions of the current pixel in each window. All correlation curves are considered as sources of information that are combined to yield the disparity of the current pixel.

B. Confidence modeling

This step aims at associating a confidence value to each correlation measure, i.e. to each cube of the correlation volume. The confidence volume has therefore the same shape as the correlation volume.

Some authors have already proposed to associate a confidence value with the matching [4], [6], [3]. In [4], the author analyses the relative difference between the two lowest minima of the correlation function. If this value is higher than a threshold, the window configuration is validated and the confidence is high. In [6], 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 compute semi-dense disparity maps [10]. As described in [5], several confidence measures of different kinds can be extracted from the correlation curves.

Since we have to combine several confidence criteria, that are not only numerical and do not share the same semantics, fuzzy logic is the most adapted framework. Fuzzy logic techniques are often used for applications in which human expertise is available to solve the given problem, and they allow for formalising fuzzy assertions in order to make them usable by a computer [11].

In our case, we compute the confidence value by: 1) extracting several criteria from the correlation curves and 2) modeling each of them by fuzzy membership functions. All the correlation curves of the correlation volume are processed by a set of fuzzy filters that yield the fuzzy confidence $Conf_{FL}(p, w_n, s)$ for every position p of the current pixel in the window of size w and for every shift s. A single confidence value is then associated to every shift in order to evaluate the likelihood that this shift is the disparity for the current pixel.

We extract the following three characteristics from the correlation curves:

- The curvature metric, which measures the curvature of the correlation curve for every shift (excluding sides). The sharper the correlation valley, the higher the curvature and the quality of the matching. The curvature metric $Cur(p, w_n, s)$ is defined as:

$$Cur(p, w_n, s) = \qquad -2 \cdot C(p, w_n, s)$$
$$+C(p, w_n, s+1) + C(p, w_n, s-1) .$$

- The rank $R(p, w_n, s)$. Correlation values are sorted in increasing order and the rank of a correlation value corresponds to its position in the sorted list. In the ideal case, a single minima exists and the shift that appears at the first position is the disparity. Unfortunately, the curve often shows several minima and the first shift in the sorted list is not always the solution. In this case, the confidence must be decreased.

- The number $N(p, w_n)$ of inflexion points characterising the number of minimum valleys (convex curves). A correlation curve with a single valley presents less ambiguity than other with several minimum valleys. $N(p, w_n)$ is defined by:

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

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

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

Figure 3 shows the membership functions of the fuzzy Curvature Metric criterion with three states: bad, medium and good.
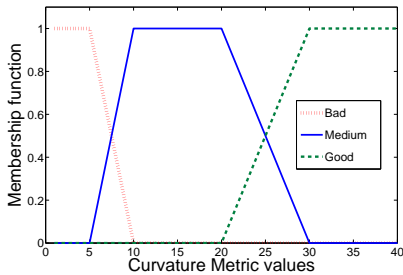

Fig. 3.   Membership functions of the fuzzy Curvature Metric criterion

In the fuzzy filters, we use the standard IF-THEN-ELSE inference mechanism. Each inference rule computes an elementary confidence $Conf$ for a given value of p, w_n and s. The 27 (3^3) inference rules have been defined by analysing the behaviour of the three characteristics, Cur, R, and N on typical image neighbourhoods. In each rule, the fuzzy confidence value $Conf$ can take five states: null, bad, medium, good, excellent.

In the defuzzification step, the 27 fuzzy confidence values are combined using the center of gravity method, to yield the global confidence value $Conf_{FL}(p, w_n, s)$. In our implementation, no parameter of this fuzzy processing is adjusted, i.e. each result is computed with an identical fuzzy logic filter with fixed parameters.

C. Decision-making proccess

In this part, we aim at defining a basic but efficient method for selecting the correct disparity in the confidence volume. Firstly, we select the shift $s^*_{w_n}(p)$ corresponding to the maximum confidence value for any position p of the current pixel in the window with size w_n:

$$s^*_{w_n}(p) = \arg[\max_s \{Conf_{FL}(p, w_n, s)\}] \ . \qquad (4)$$

Then, we compute the number $N_{w_n}(s)$ of occurrences of these maximum values, for a fixed width w_n, as:

$$N_{w_n}(s) = Count_p[s = s^*_{w_n}(p)] \ , \qquad (5)$$

and we mark the disparity $d^*_{w_n}$ at the shift value s with the maximum number of occurrences for all positions of the current pixel:

$$d^*_{w_n} = \arg[\max_s(N_{w_n}(s))] \ . \qquad (6)$$

Finally, the disparity $d^*_{w_n}$ is assigned to the pixel (x, y) if the following condition is satisfied:

$$\frac{N_{w_n}(d^*_{w_n})}{2n+1} \geq T_{dec} \ . \qquad (7)$$

where T_{dec} is a threshold. When T_{dec} is equal to one, the condition of equation (7) means that the same disparity was computed for all the positions of the reference pixel in the window with size w_n.

The disparity map is filled by repeating the algorithm steps equivalent to equations (5), (6) and (7), starting with the largest window and decreasing its width until a disparity is assigned to the pixel. This process allows for affecting the disparity in untextured areas and for filling the map near discontinuities while avoiding errors (essentially due to larger correlation windows).

To sum up, the decision-making step computes the disparity at each pixel (x, y) as follows:

$$d(x, y) = 0 \ ; \ n = n_{max}$$
Do
$$s^*_{w_n}(p) = \arg[\max_s \{Conf_{FL}(p, w_n, s)\}]$$
$$N_{w_n}(s) = Count_p[s = s^*_{w_n}(p)]$$
$$d^*_{w_n} = \arg[\max_s(N_{w_n}(s))]$$
If $N_{w_n}(d^*_{w_n}) \geq T_{dec} \times (2n+1)$
$$d(x, y) = d^*_{w_n}$$
Else $n - 1$
While $(n > 0)$ and $(d(x, y) = 0)$

The final confidence value $Conf(x, y)$ associated with the selected disparity $d(x, y)$ is computed as the average of the confidence values of the selected shift, for each position of the reference pixel in a window with the selected width.

Thus, the confidence volume is analysed to reduce the amount of data and to provide a confidence for every shift. It is important to note that several levels of decision can be implemented using this approach. The standard disparity determination corresponds to a basic decision that selects a single shift for each pixel and associates with it a confidence varying from 0 to 100%.

This method yields a disparity map associated to its confidence. The following section presents results obtained on the $Tsukuba$ stereo image pair in grey levels.

III. RESULTS ON GREY LEVEL IMAGE

This section, presents and compares results based on our 1D method and those based on the classical 2D Hirschmüller method [4]. Hirschmüller combines the correlation computed on the window centered on the considered pixel with the correlations computed on several support windows [4]. 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.

We have evaluated our stereovision technique on the $Tsukuba$ stereo image pair. Figure 4 presents the left image of the $Tsukuba$ stereo pair, the ground truth, the textured and untextured regions and the areas near discontinuities.

Our 1D method yields a quasi-dense disparity map associated with a confidence map as presented figure 5. In this comparison, the threshold T_{dec} was fixed to one and the disparity maps are sparsely defined.
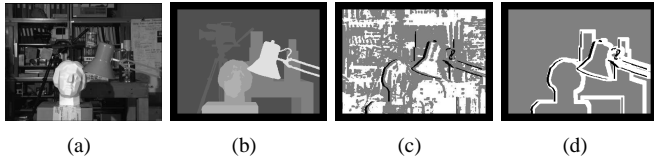

Fig. 4. (a) Left image, (b) Ground truth disparity map, (c) Textured regions (grey pixels) and untextured regions (white pixels), (d) Occluded regions (black pixels) and regions with discontinuities (white pixels).

Figure 5 presents the left image in grey levels of the $Tsukuba$ pair, the ground truth, the quasi-dense disparity map computed and the associated confidence map.
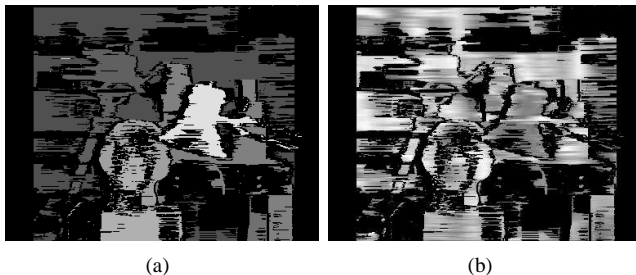

Fig. 5. (a) Quasi-dense disparity map computed on grey level image, (b) Associated confidence map.

TABLE I

RESULTS OBTAINED ON GREY LEVEL IMAGES BASED ON MIDDLEBURY
STEREO EVALUATION WITH $|\Delta\varepsilon| > 0.5$

Algorithm	Tsukuba			
	$B_{\overline{\mathcal{O}}}$	$B_{\overline{\mathcal{T}}}$	$B_{\mathcal{D}}$	$B_{\mathcal{T}}$
Density	45.15	23.11	46.45	62.94
Realtime [4]	3.67^2	2.89^2	4.32^2	4.24^2
Our method	2.19^1	1.24^1	3.03^1	2.88^1

For the quantitative comparison, we compute the pixel to pixel difference between our disparity map and the ground truth disparity map of the stereo pair. Because it is a quasi-dense disparity map, we do not take into account the non-assigned pixels. To start with an identical basis of comparison, we proceed in the same way for the other method, i.e. taking only into account the pixels to which a disparity was assigned by our method. We consider that a disparity is false if the pixel to pixel difference is greater than 0.5, which

is the lowest error threshold defined in the new evaluation technique of the Middlebury website [12].

Table I presents the results of this quantitative comparison. The first row gives the density of the computed disparity map and the following rows the error rates for each algorithm versus the type of region: non-occluded ($B_{\overline{\mathcal{O}}}$), texture-less ($B_{\overline{\mathcal{T}}}$), near discontinuities ($B_{\mathcal{D}}$) and textured ($B_{\mathcal{T}}$).

Firstly, one can notice that our method is top ranked. Every object in this image is located in a different vertical plane, which does not allow an error threshold greater than 0.5. Thus, the averaging effect of 2D correlation windows yields a lot of matching errors.

With our proposed technique, we consider that in untextured regions, there is no information and that it is more appropriate to assign no disparity to the pixel. This behaviour can be observed in the resulting map where the density of matched pixels is very low in untextured areas.

In regions near discontinuities, we achieve good results because objects are located in vertical planes distant from each other. 2D method violates the fronto-parallel assumption.

Therefore, a 2D correlation window including several raster lines does not satisfy the constraint of constant disparity along the vertical direction and averages the information. In the next part, we introduce colour information in our 1D method in order to show its capabilities to improve matching results.

IV. USE OF COLOUR INFORMATION

In this section, we aim at using colour information available in the $Tsukuba$ image pairs; The initial colour space of the image (i.e. the RGB colour space) is retained as basis of this first study.

A. Combination method

The RGB colour space can be decomposed into its three components: red, green and blue. Our 1D method is applied on each component and yields three disparity maps with the confidence maps. At this stage, we ought to combine the three disparity maps in order to obtain a final disparity map. The following basic rule is employed to determine the final disparity value: For a given pixel (x, y), if the assigned disparity values on the different colour components are identical, this value is retained as the final disparity. The associated confidence value is computed as the average of the confidence values of the assigned disparities. If one or several assigned disparities are different, the value of pixel (x, y), in the disparity and confidence maps, is equal to zero.

The following section presents a comparison between the disparity maps computed with the two ways : the grey level one and the colour one.

B. Results on colour image and comparison

The previously described combination method allows merging the disparity values obtained with the three components of the RGB colour space in order to compute the final disparity and confidence maps.

The obtained results are shown figure 7. Like in the previous comparison, the threshold T_{dec} was fixed to one.

We can notice that the disparity maps computed on the three components (red, green and blue) are as dense as the one obtained on grey level image. However, each component does not assign the same pixels. For example, we can see that the orange lamp, composed essentially of the red component, is totally affected on the disparity map of the red component; the green and blue ones are sparser. Thus, by combining the disparity maps obtained with the three components, our 1D method computes denser disparity map than with grey level image.

In a quantitative way, we compute error rates based on Middlebury Stereo Evaluation (with $|\Delta\varepsilon| > 0.5$) and this, on the same basis of assigned pixels. In table II, we can notice that the disparity values computed with the colour way are slightly more precise, especially in areas near discontinuities.

TABLE II

ERROR RATES OF COMMON ASSIGNED DISPARITY VALUES BASED ON MIDDLEBURY STEREO EVALUATION WITH $|\Delta\varepsilon| > 0.5$

Method	Tsukuba			
	$B_{\overline{\mathcal{O}}}$	$B_{\overline{\mathcal{T}}}$	$B_{\mathcal{D}}$	$B_{\mathcal{T}}$
Grey levels	1.89	0.92	2.66	2.60
RGB-Fusion	1.88	0.92	2.58	2.57
Density	44.15	21.99	45.15	62.02

By tunning the confidence value between 0 and 100%, we can remove assigned disparity values which seem less confident. In this way, the density of the disparity map varies with the selected confidence level: This allows determining the evolution of the matching error rates versus the density of the map. Figures 6(a), 6(b), 6(c) and 6(d) show matching error rates versus density rates based on the disparity maps computed with grey level images and RGB images in respectively non-occluded areas ($B_{\overline{\mathcal{O}}}$), non-textured areas ($B_{\overline{\mathcal{T}}}$), textured areas ($B_{\mathcal{T}}$), and discontinuities ($B_{\mathcal{D}}$). It is obvious that the colour approach of our 1D method provides, for identical density rates, better matching results that the grey level approach. Moreover, by combining the disparity obtained with the three RGB components, we can obtain denser disparity maps with more precision.

In the next section, we compare our 1D approach to colour correlation-based stereo matching with other similar methods.

C. Comparison with similar methods

In this section, we compare our results with those obtained by other similar algorithms. We focus our attention on window-based stereo matching algorithms associated to the WTA approach and using the colour information, for which associated results are available in [12]. We have selected four comparable methods:

- MMHM [6] ;
- Improved Coop. [13] ;
- Adapt. weights [14] ;
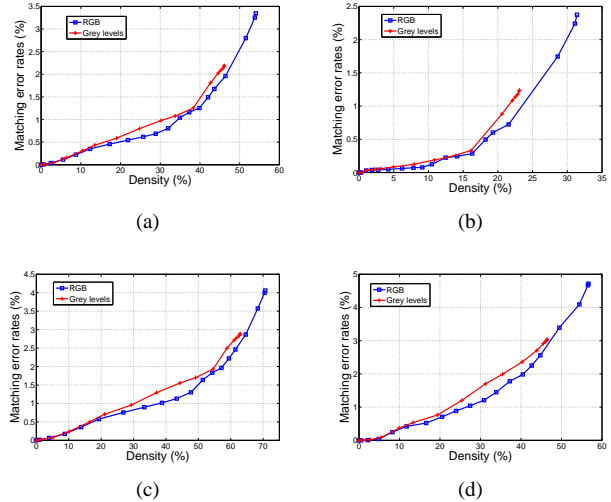- Comp. win. [15] ;

Fig. 6. Matching error rates versus density rates based on the disparity maps computed with grey level images and RGB images in: (a) non-occluded areas ($B_{\overline{\mathcal{O}}}$), (b) non-textured areas ($B_{\overline{\mathcal{T}}}$), (c) textured areas ($B_{\mathcal{T}}$), and (d) discontinuities ($B_{\mathcal{D}}$)

The following results, presented in Table III, were obtained from $Tsukuba$ stereo pair using the evaluation technique proposed by the Middlebury website. They are compared with the results provided on this website for the four previously presented algorithms. The first row gives the density of the computed disparity map and the following rows the error rates for each algorithm versus the type of region: non-occluded ($B_{\overline{\mathcal{O}}}$), texture-less ($B_{\overline{\mathcal{T}}}$), near discontinuities ($B_{\mathcal{D}}$) and textured ($B_{\mathcal{T}}$).

Firstly, one can notice that our method gives results comparable with those of others algorithms although it uses only 1D correlation windows: Our method is ranked in the first places. Firstly, this is due to our 1D approach which provides more precise matching in comparison with 2D approaches. Moreover, the combination of the three precise disparity maps, obtained on each color component, computes a final disparity map denser and always also precise.

Our method presents very interesting results especially in regions near discontinuities: The error rates obtained with the four other methods are much higher. In the same way that with the grey level method, the untextured regions are very sparse on the disparity map because no information allows to have efficient matching, even if the colour information is used.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed an original stereo matching method based on the analysis of a set of 1D correlation scores. The method computes a quasi-dense disparity map and an associated confidence map.

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

TABLE III

COMPARISON BETWEEN OUR RESULTS OBTAINED ON RGB COLOUR
IMAGES AND FOUR SIMILAR COLOUR METHODS (BASED ON
MIDDLEBURY STEREO EVALUATION WITH $|\Delta\varepsilon| > 0.5$)

Method	$B_{\overline{\mathcal{O}}}$	$B_{\overline{\mathcal{T}}}$	$B_{\mathcal{D}}$	$B_{\mathcal{T}}$
Density	53.08	31.40	56.66	70.67
Improved Coop. [13]	4.38[3]	3.65[3]	5.79[3]	4.91[3]
Adapt. Weights [14]	3.25[1]	1.41[1]	5.77[2]	4.60[2]
Comp. Win. [15]	4.95[4]	3.71[4]	6.48[4]	5.85[5]
MMHM colours [6]	4.99[5]	4.72[5]	8.68[5]	5.19[4]
Our 1D colour method	3.35[2]	2.37[2]	4.72[1]	4.06[1]

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.

Our method has been applied on the well-known stereo image pair of the university of $Tsukuba$ in grey levels : it shows the ability to provide good matching in most cases. The algorithm takes the advantage of the 1D property and yields much better results than the classical 2D method presented by Hirschmüller.

So, the algorithm has been applied on each colour component of the RGB colour space ; then, the three disparity maps, obtained on each colour component, are combined to compute the final disparity map with its confidence map. Our colour method has been compared to several similar colour methods presented on the Middlebury stereo website with the $Tsukuba$ stereo pair.

This study has shown that the information in a 1D window is often rich enough to provide good matching results. Moreover, the information brought by each colour component allows to have denser disparity map in comparison of those obtained on grey level image.

The advantages of our method 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, we study how to analyse in a different way the correlation and confidence volumes, in order to reduce the disparity error rate. In the same way, we aim at computing disparity maps on other colour spaces and find the more appropriated space with our algorithm and the studied image pair. So, disparity maps could be computed on colour components of a set of colour spaces: a decision step could find the most confident component among the whole of studied components.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] 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.

[2] 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.

[3] 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.

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

[5] 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.

[6] 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.

[7] A. Fusiello, E. Trucco, and A. Verri, "Rectification with unconstrained stereo geometry," Essex, U.K., Sept. 1997, pp. 400–409.

[8] N. Vandenbroucke, L. Macaire, and J. Postaire, "Color image segmentation by pixel classification in an adapted hybrid color space. application to soccer image analysis," *Computer Vision and Image Understanding*, vol. 90, no. 2, pp. 190–216, 2003.

[9] S. Chambon and A. Crouzil, "Colour correlation-based matching," *International Journal of Robotics and Automation*, vol. 20, no. 2, pp. 78–87, 2005.

[10] R. Šára, "Finding the largest unambiguous component of stereo matching," in *European Conference on Computer Vision*, O. Berlin, Germany: Springer, May 2002, p. 919.

[11] L. Zadeh, "Fuzzy sets," in *Information and Control*, vol. 8, June 1965, pp. 338–353.

[12] D. Scharstein and R. Szeliski, "A taxomomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1, pp. 7–42, Apr. 2002, the evaluation page URL: **www.middlebury.edu/stereo**.

[13] H. Mayer, "Analysis of means to improve cooperative disparity estimation," in *ISPRS Conference on Photogrammetric Image Analysis*, Sept. 2003, pp. 25–31.

[14] K. Yoon and I. Kweon, "Locally adaptive support-weight approach for visual correspondence search," in *IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, San Diego, CA, USA, June 2005, pp. 924–931.

[15] O. Veksler, "Stereo matching by compact windows via minimum ratio cycle," in *International Conference on Computer Vision*, Vancouver, Canada, July 2001, pp. 540–547.

Fig. 7. (a) Left image, (b) Quasi-dense disparity map, (c) Associated confidence map.