2011 14th International IEEE Conference on
Intelligent Transportation Systems
Washington, DC, USA. October 5-7, 2011

A Real-time Spatio-Temporal Stereo Matching for Road Applications

Mohamed El Ansari, Abdenbi Mazoul, Abdelaziz Bensrhair, George Bebis

Abstract— **This paper presents a real-time approach for matching stereo images acquired by a stereo sensor embedded in a moving vehicle. The new method consists of matching edge points extracted from stereo images using the temporal relationship, which exists between consecutive stereo pairs. Matching a current stereo pair takes into account the matching results of the preceding stereo pair. The method looks first for what we call *matching control edge points* (MCEPs) based on spatio-temporal matching of edge curves of consecutive stereo pairs. Dynamic programming is considered for matching edge points of the stereo images. The MCEPs drive the optimal path of the dynamic programming. The proposed approach has been tested on virtual and real stereo image sequences and the results are satisfactory.**

I. INTRODUCTION

An intelligent vehicle (IV) can perform road obstacle detection by knowing its environment. Stereo vision [1] is a well-known method used to obtain an accurate and detailed 3D representation of the environment around an IV. The key problem in stereo vision consists of finding correspondences between pixels of stereo images taken from different viewpoints [2]. Exhaustive surveys on methods tackling the correspondence problem are available in [3], [4]. A taxonomy of dense stereo correspondence algorithms together with a testbed for quantitative evaluation of stereo algorithms is provided by Scharstein and Szeliski [5]. The taxonomy shows that graph cuts-based methods [6] outperform other methods, but they are time consuming which makes them not suitable for real-time applications (e.g. advanced driver assistance systems (ADAS)).

Although there is strong support that the incorporation of temporal information can achieve better results [7], [8], [9], [10], only a small amount of research has been devoted to the reconstruction of dynamic scenes from stereo image sequences. We believe that by considering the temporal consistency between consecutive frames, stereo matching results could be improved [11], [12]. This paper presents a new stereo matching method exploiting the connection that exists between consecutive stereo pairs of images provided by a stereo cameras aboard a vehicle. We use the same approach detailed in [11] to find such a connection. Based on this connection, we first search for spatio-temporal matching of significant edge curves in consecutive stereo images. This allows to find correspondences between edge curves in the current stereo pair. Second, the correspondences between edges points of the matched edge curves is deduced. Third, the matched edge points are used to drive dynamic programming [13] search for matching all edge points in the current stereo pair. The proposed method has been tested both on virtual and real stereo image sequences and showing promising results.

The remainder of the paper is organized as follows. Section II overviews stereo methods handling stereo sequences and using temporal consistency. The new stereo method is detailed in section III. Experimental results are presented in section IV. Section V concludes the paper.

M. El Ansari is with the LabSIV, Department of Computer Science, Faculty of Science, University of Ibn Zohr, 80000 Agadir, Morocco, and the LITIS Lab., INSA-Rouen, France. m.elansari@univ-ibnzohr.ac.ma, melansari@gmail.com

A. Mazoul is with the LabSIV, Department of Computer Science, Faculty of Science, University of Ibn Zohr, 80000 Agadir, Morocco. mazoul_abdenbi@hotmail.com

A. Bensrhair is with the LITIS EA 4108, INSA de Rouen, Avenue de l'Université, BP 8, 76801 Saint-Etienne-du-Rouvray Cedex, France. Abdelaziz.Bensrhair@insa-rouen.fr

G. Bebis is with the CVL Lab., Department of Computer Science and Engineering, University of Nevada, Reno, NV 89557, USA, and the Computer Science Department, King Saud University, Riyadh, Saudi Arabia. bebis@cse.unr.edu

II. RELATED WORK

In recent years, several techniques have been proposed to obtain more accurate disparity maps from stereo sequences by utilizing temporal consistency [7], [9], [14], [8]. Most of these methods use either optical flow or spatiotemporal window for matching stereo sequences. In their approach, Tao et al. [14] proposed a dynamic depth recovery where a scene representation, that consists of piecewise planar surface patches, is estimated incrementaly. Such a representation is derived based on color segmentation. Each segment is modeled as a 3D plane. The motion of the plane is described using a constant velocity mode. A spatial match measure and a scene flow constraint [15] are employed in the matching process. The accuracy of the results and the processing speed are limited by the image segmentation algorithm used. Vedula et al. [15] presented a linear algorithm to compute 3D scene flow based on 2D optical flow and estimate 3D structures from the scene flow. In [16], the temporal consistency was enforced by minimizing the difference between the disparity maps of adjacent frames. This approach is designed for offline processing only (i.e. it takes pre-captured stereo sequences as input and calculates the disparity maps for all frames at the same time). In [9], an algorithm was developed to compute both disparity maps and disparity flow maps in an integrated process. The disparity map generated for the current frame is used to predict the disparity map for the next frame. The disparity map found provides spatial correspondence information which was used to cross-validate the disparity flow maps estimated for different views.

978-1-4577-2196-0/11/$26.00 ©2011 IEEE

Programmable graphics hardware were used for accelerating processing speed.

Zhang et al. [8], proposed to extend existing traditional methods by using both spatial and temporal variations. The spatial window used to compute SSD (sum of squared differences) cost function is extended to a spatiotemporal window for computing the sum of SSD (SSSD). Their method could improve results when we deal with static scenes and structured light. It fails to do so with dynamic scenes. Davis et al. [7] developed a similar framework as in [8]. However, their work is focused on geometrically static scenes imaged under varying illumination. Given an input sequence taken by a freely moving camera, Zhang et al. [17] proposed a novel approach to construct a view-dependent depth map for each frame. Their method takes one sequence as input and provides the depth for different frames (i.e. offline processing). It is not applicable in an IV.

Recently, the authors proposed the so-called association as a method to use for finding the relationship between consecutive stereo pairs [11]. In this paper, we propose to use the same principle as in [11] to get the connection between adjacent (consecutive) frames. A spatio-temporal approach is presented to match edge curves of adjacent frames. The edge points of matched edge curves will be used in the matching process to drive dynamic programming.

III. Stereo matching algorithm

This section describes the steps of the proposed method for matching stereo images captured by stereo sensor mounted aboard an IV. We note that the stereoscopic sensor used in our experiments provides rectified images (i.e., corresponding pixels have the same *y-coordinate*). The principle of the new approach is to exploit the link between consecutive stereo pairs. So, the matching results of the preceding stereo pair are used for matching the current stereo pair. The following notations will be used in the rest of the paper. I_{k-1}^L and I_{k-1}^R denote the left and right stereo images of the frame f_{k-1} acquired at time $k-1$ and d_{k-1} is the corresponding disparity map. I_k^L and I_k^R represent the left and right stereo images of the frame f_k acquired at time k. We assume that the frame $f_k = (I_k^L, I_k^R)$ represents the current stereo pair for which we want to compute the disparity map d_k. $f_{k-1} = (I_{k-1}^L, I_{k-1}^R)$ represents the preceding frame for which the disparity map d_{k-1} is available. The matching problem is formulated as follows: how to compute d_k by taking into account d_{k-1} and the relationship between the frames f_{k-1} and f_k.

A. Edge detection

The first step consists of extracting significant features from the stereo images to be matched. In this work, we are interested in using edge points for matching. The Canny edge detector [18] is regarded as one of the best edge detectors currently in use. It provides continuous edge curves, which are vital to the proposed matching method. Consequently, we use the Canny operator for edge (points and curves) detection from the stereo images. In the rest of the paper we use the following notations. $S_f^m = \left\{ C_f^{m,i} \right\}_{i=1,..,N_f^m}$ denotes the set of edge curves extracted from the image I_f. $f \in \{k, k+1\}$ represents the frame index. $m \in \{L, R\}$ is the index of the stereo image, i.e. L for left image and R for right image. N_f^m represents the number of edge curves in the image I_f^m.

In [11], the declivity operator [**?**] was used to detect edge points. This operator does not detect horizontal edge curves. Using the Canny detector in the current work allows the detection of more edge points, which results in more dense disparity maps.

B. Association between edge points of consecutive images

As mentioned earlier, the main idea of the proposed approach is to exploit the relationship between consecutive stereo pairs. The authors have proposed the so-called *association* to achieve this task [11]. This subsection describes the method used to find the association between edge points of consecutive frames (i.e., association between edge points of the images I_{k-1}^L and I_k^L (resp. I_{k-1}^R and I_k^R)).

Let us assume that we want to find the association between the left images. Let us consider two edge points P_{k-1}^L and Q_{k-1}^L belonging to a curve $C_{k-1}^{L,i}$ in the image I_{k-1}^L and their corresponding ones P_k^L and Q_k^L belonging to a curve $C_k^{L,j}$ in the image I_k^L (see Fig. 1). The associate point to point P_{k-1}^L is defined [11], [12] as the point belonging to the curve $C_k^{L,i}$ whith the same *y-coordinate* as P_{k-1}^L. Two associate points are two edge points belonging to two corresponding curves of two consecutive images of the same sequence and having the same y-coordinate. From Fig. 1, we remark that point Q_k^L constitutes the associate point of P_{k-1}^L. More details about how to find the association between consecutive images is available in [11].

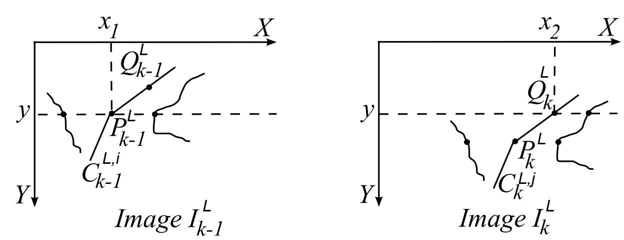

Fig. 1. I_{k-1}^L and I_k^L represent consecutive images of the left sequence. The point Q_k^L in the image I_k^L constitutes the associate point of the point P_{k-1}^L in the image I_{k-1}^L.

For each edge point in image I_{k-1}^L (resp. I_{k-1}^R) we look for its associate one, if it exists, the image I_k^L (resp. I_k^R).

C. Spatio-temporal matching of edge curves of consecutive stereo images

Here, we illustrate how to find the correspondence between edge curves of consecutive frames f_{k-1} and f_k based on the association computed in the previous subsection and the known disparity map d_{k-1} of the frame f_{k-1}.

1) Temporal correspondence: This involves matching between curves of the image I_{k-1}^L (resp. I_{k-1}^R) with edge curves of the image I_k^L (resp. I_k^R). We illustrate the proposed method using the images I_{k-1}^L and I_k^L. The same process is used to match the edge curves between the images I^R_{k-1} and I^R_k.

Temporal correspondence consists of finding for each edge curve $C^{L,i}_{k-1}$ in the set S^L_{k-1} its corresponding edge curve $C^{L,j}_k$ in the set S^L_k, if it exists. Let $Ass(C^{L,i}_{k-1}) = \{ae_n\}_{n=1,..,N_i}$ be the set of edge points ae_n, belonging to the image I^L_k, which represent the associates of the edge points of the edge curve $C^{L,i}_{k-1}$. N_i is the number of associations found for the edge curve $C^{L,i}_{k-1}$. If M_i represents the number of edge points in $C^{L,i}_{k-1}$, $N_i \leq M_i$ because there are edge points in image I^L_{k-1} for which there is no associate in image I^L_k. If there is no error in the association process, all the edge points belonging to the set $Ass(C^{L,i}_{k-1})$ should belong to one edge curve, which is the corresponding curve to $C^{L,i}_{k-1}$. Unfortunately, there are some errors inherent to the association process. Consequently, the edge points ae_m may belong to different curves in S^L_k. We find the match of $C^{L,i}_{k-1}$ by looking for the curve $C^{L,j}_k$, which contains the maximum number of edge points in $Ass(C^{L,i}_{k-1})$. We apply the same method to all the edge curves in S^L_{k-1} to find their corresponding ones in S^L_k.

2) Spatial correspondence: This step involves matching the edge curves of the stereo images I^L_{k-1} and I^R_{k-1} on the basis of the disparity map d_{k-1}. The same principle as in temporal correspondence is used to find spatial correspondence.

Let $Match(C^{L,i}_{k-1}) = \{me_n\}_{n=1,..,N_i}$ be the set of edge points me_n, belonging to the image I^R_{k-1}, which match the edge points of $C^{L,i}_{k-1}$. N_i is the number of matched edge points belonging to $C^{L,i}_{k-1}$. If M_i represents the number of edge points in $C^{L,i}_{k-1}$, $N_i \leq M_i$ because there is a number of edge points in the image I^L_{k-1} for which there is no match in the image I^R_{k-1}. If there is no error in the matching process, all the edge points belonging to the set $Match(C^{L,i}_{k-1})$ should belong to one edge curve, which is the corresponding of the curve $C^{L,i}_{k-1}$. Unfortunately, there are some errors inherent to the matching process. Consequently, the edge points me_m may belong to different curves in S^R_{k-1}. We find the match of the curve $C^{L,i}_{k-1}$ by looking for the curve $C^{R,j}_{k-1}$, which contains the maximum number of edge points in $Match(C^{L,i}_{k-1})$.

At this point, we have (1) the spatial correspondence between the edge curves of the stereo images I^L_{k-1} and I^R_{k-1}, (2) the temporal correspondence between the edge curves of the images I^L_{k-1} and I^L_k, and (3) the temporal correspondence between the edge curves of the images I^R_{k-1} and I^R_k. We can deduce easily the spatial correspondence between the edge curves of the images I^L_k and I^R_k as depicted in Fig. 2. Let $C^{L,i}_k$ be an edge curve in I^L_k. Finding the spatial match of $C^{L,i}_k$ is achieved in four steps. The first step searches the match of $C^{L,i}_k$ in I^L_{k-1}, which we call $C^{L,j}_{k-1}$. In the second step, the corresponding edge curve, $C^{R,m}_{k-1}$, of $C^{L,j}_{k-1}$ is searched in the image I^R_{k-1}. The third step looks for the match, $C^{R,n}_k$, of $C^{R,m}_{k-1}$ in the image I^R_k. The last step deduces that $C^{R,n}_k$ represents the match of $C^{L,i}_k$. We repeat the same process for all the edge curves of the image I^L_k to find their matches in the image I^R_k.

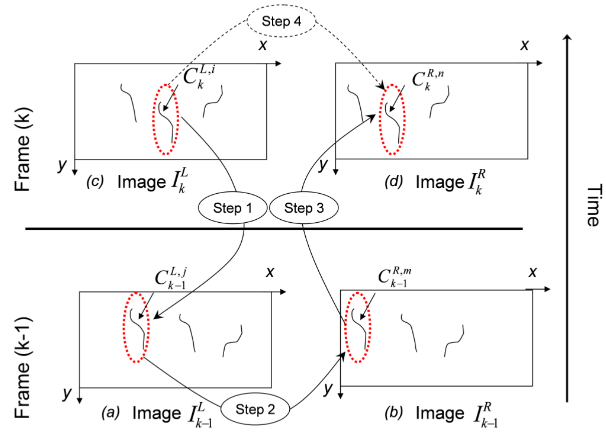

Fig. 2. Spatial and temporal matching of the edge curves belonging to consecutive frames.

D. Initial disparity map

Here, we illustrate how to get an initial disparity map for the stereo pair f_k on the basis of matched pairs of edge curves of the same frame. Let $C^{L,i}_k$ and $C^{R,n}_k$ be corresponding edge curves belonging to the images I^L_k and I^R_k, respectively, and e^L_m an edge point belonging to $C^{L,i}_k$. Known that the stereo images are rectified, the corresponding edge points should have the same y-coordinate. The match of e^L_m, if it exists, should belongs to $C^{R,n}_k$ and have the same y-coordinate as e^L_m. We repeat this process for all edge points of $C^{L,i}_k$ to find their matches in $C^{R,n}_k$. The same method will be applied to all pairs of matched edge curves of the frame f_k. As a result, we get a number of pairs of matched edge points in the stereo images of the frame f_k. These correspondences allow to generate what we call an initial disparity map (IDM) for the current frame. The IDM is more accurate compared to the so-called pre-estimated disparity map computed in [11]. We refer to the matched edge points as **MCEPs**. They will be used to drive the dynamic programming for matching the remaining edge points of the current frame (section III-E.3).

E. Stereo matching method of edge points of the current frame

In section III-D, we described how to match the edge points belonging to the edge curves of the frame f_k. This section presents the method we propose for matching the remaining edge points of the current frame by considering MCEPs.

1) Disparity range constraint: The accurate choice of the maximum disparity threshold value is crucial to the quality of the output disparity map and computation time [5], [19]. In [11], [12], the authors presented a method to compute possible disparities (disparity range). The method is based on analyzing the v-disparity [20] computed from IDM. This provides the disparity range for each scanline of the stereo images. We use the same idea here to determine the disparity range for each image line on the matched stereo pair. More details can be found in [11].

2) Cost function: As a similarity criterion between corresponding edge points, we use a cost function based on the gradient magnitude and orientation at the matched edge points. Let e^L and e^R be two edge points belonging to images I_k^L and I_k^R, respectively. We denote by m^L and m^R (resp. θ^L and θ^R) their gradient magnitudes (resp. orientations), respectively. We assume that corresponding edge points on the stereo images should have the same (or closer) gradient magnitudes as well as the same (or close) orientations. Therefore, we define the cost function as follows.

$$C(e^L, e^R) = \Big\{ \big(I_k^L(x^L, y^L) - I_k^R(x^R, y^R)\big)^2 + (m^L)^2 + (m^R)^2 - 2 * m^L * m^R * cos(\theta^L - \theta^R) \Big\}^{1/2} \tag{1}$$

where (x^L, y^L) and (x^R, y^R) are the coordinates of the edge points e^L and e^R, respectively.

3) Dynamic programming: Let $\{e^L_{i,sl}\}_{i=1,..,N^L_{sl}}$ (resp. $\{e^R_{j,sl}\}_{j=1,..,N^R_{sl}}$) be the set of edge points in the scanline (image line) sl of the image I_k^L (resp. I_k^R), which are ordered according to their x-coordinates, where N^L_{sl} (resp. N^R_{sl}) is their number. We demonstrate how to match these edge points for the scanline sl using dynamic programming. The same technique will be used for all scanlines of the stereo images I_k^L and I_k^R.

The problem of obtaining correspondences between edge points on right and left epipolar scanlines can be solved as a path finding problem on the 2D plane [13]. We propose to subdivide the search space into a number of sub-spaces depending on the number of MCEPs found at the scanline sl.

Fig. 3 illustrates an example of 2D search plane, which is divided into sub-search planes. The vertical lines show the positions of edge points on the left scanline and the horizontal ones show those on the right scanline. We refer to the intersections of those lines as nodes. Nodes in this plane correspond to the stages in dynamic programming where a decision should be made to select an optimal path to that node. They represent the candidate matches. Optimal matches are obtained by the selection of the path, which corresponds to a minimum value of the global cost. For each sub-search plane, the optimal path must goes from the upper left corner to the lower right corner monotonically due to the condition on ordering. Because of the non reversal ordering constraint, starting from a node in the search plane, a path can be extended towards only one of the three directions: east, south, or southeast.

As depicted in Fig. 3, each sub-search plane consists of matching the edge points between consecutive MCEPs. As an example the second sub-search plane #2 is used to match the edge points situated between e_1^L and e_2^L in the left scanline sl of the image I_k^L with the edge points located between e_1^R and e_2^R in the right scanline sl of the image I_k^R. First, the disparity range is used to select the valid nodes for the sub-search plane #2. Second, the cost function (Eq. 1) is used to fill in the valid nodes. After looking for the optimal path, the pairs of corresponding edge points between e_1^L and e_2^L in the left scanline and those between e_1^R and e_2^R in the right scanline are determined. The same process is repeated for all the sub-search planes. The same method is applied to all other scanlines for matching the edge points of the whole image.

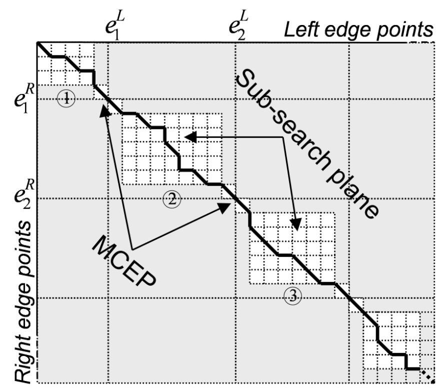

Fig. 3. 2D search plane subdivided into sub-search planes. The horizontal axis corresponds to the left scanline and the vertical one corresponds to the right scanline. Vertical and horizontal lines are the edge points positions and path selection is performed at their intersections.

IV. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed approach, we have experimented with virtual and real stereo sequences. Also, we have compared it to the method presented in [11] to assess its performance.

First, we used the MARS/PRESCAN virtual stereo images available in [21]. The size of the images is 512×512. The left stereo image of the frame #293 of the virtual stereo sequences is shown on the left side of Fig. 4. The edge image, obtained by the canny detector, is depicted in the right side of Fig. 4. The corresponding disparity map computed by the new approach is shown in Fig. 5. We have used false colors for representing the disparity map. Table I summarizes the matching results obtained with the new method and the one in [11]. It shows the number of matched edge points (NME) and the percentage of correct matches (PCM) for the frame #293. It is clear that the new method gives more correct matches. The percentage of false matches with the new method is less than the other one. Table I justifies clearly the performance of the new method. The same performance was obtained on the other frames of the virtual sequence.

The proposed method has also been tested on the real sequence depicted in Fig. 6. The image size is 384×288. The stereo sequence was acquired by a stereo vision sensor embedded in a moving car. The velocity of the car was 90km per hour. The stereo vision sensor provides 10 frames per second. The extracted edge points are shown in the right image of Fig. 6. The disparity map computed by the new method is illustrated in Fig. 7.

Fig. 4. (left) Left stereo image #293 of the virtual stereo sequences and (right) the edge image extracted by the canny detector.

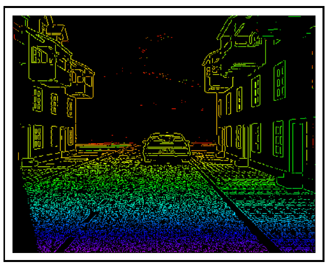

Fig. 5. The disparity map computed with the new method for the frame #293.

Method	NME	PCM
Method in [11]	15934	88.03
New method	37237	94.43

TABLE I

SUMMARY OF THE RESULTS OBTAINED BY THE NEW METHOD AND THE OTHER PRESENTED IN [11] WHEN APPLIED TO THE VIRTUAL SEQUENCE.

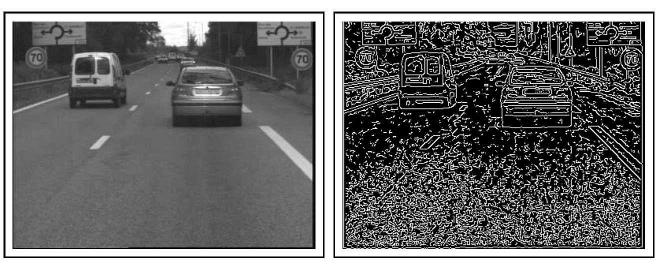

Fig. 6. (left) Left image #4180 of the real stereo sequences and (right) the corresponding edge image obtained by the canny operator.

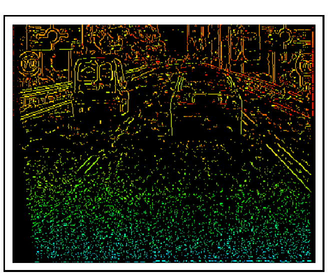

Fig. 7. Disparity map computed by the proposed method.

According to the disparity smoothness constraint, the edge points belonging to the same contour (object) should have very close or similar disparity values. So, by focusing our attention to the disparity map, we can observe that the disparity values are homogeneous at the two cars. If we focus our attention to the disparity values at the edge curves of the left and right sides of the right car, we observe that the two sides of the car have yellow color, which corresponds to the true value of the disparity. We have applied the method in [11] to the real sequence. After manually analyzing the disparities at the edge curves of the two sides of the right car, we found that the new method yields more pairs of correct matches. With the new method (resp. the method [11]), we obtained 140 (resp. 110) pairs of correct matches.

The hardware used for the experiments is a HP Intel(R) Core(TM)2 Duo CPU 2.09GHZ running under Windows XP. The running time is less than 0.3 seconds per frame.

Unfortunately, there is no public real stereo sequences available to test on our algorithm. Our method has been applied to different real stereo sequences, which we get from VisLab, University of Parma, Italy. We included only the results obtained on the stereo sequence shown in Fig. 6 because of the limited space allowed to the publication of the paper. In our knowledge, the spatio-temporal related bibliography is not exhaustive, which justifies the reason we have compared the new method with the other method proposed with the same authors in [11].

The performance of the new method can be justified as follows:

- Instead of the declivity operator used in [11], the current approach uses the Canny operator to detect edge points. This helps getting more edge points in the stereo images as the declivity operator is not able to detect horizontal edge curves. Therefore, the STM method provides more matched pairs of edge points.
- The IDM is a crucial component in both STM and TCM methods. The disparity range is derived from IDM for both STM and TCM approaches. In [11], IDM is computed from the association between edge points of the consecutive images together with the correspondences between the edge points of the preceding frame. However, in the current paper, IDM is derived from the spatio-temporal matching of the the edge curves of the consecutive frames. Consequently, the IDM computed by the new method is more accurate than the one computed by [11].
- In [11], one search space is used to find correspondences between edge points of corresponding scanlines. In the current paper, MCEPs are used to divide the search space into a number of subspaces. This enforces the dynamic programming search path to cross the pairs of matched edge points deduced from IDM. Therefore, the matching results found by the new method are more improved compared to those obtained by the method presented in [11].

V. Conclusion

In this paper, we have presented a real-time stereo matching method devoted to road applications. The new method is based on the temporal consistency between consecutive stereo pairs of images. It uses dynamic programming for matching edge points of stereo images. The dynamic programming is driven by the MCEPs, which are derived from the spatio-temporal matching of edge curves of consecutive frames. The proposed approach has been tested on different stereo sequences and the results obtained are promising. It is very fast, which makes it suitable to real-time road applications. For future work, we wish to use the matching results for obstacle detection and tracking.

References

[1] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based intelligent vehicles: State of the art and perspectives," *Robotics and Autonomous Systems*, vol. 32, pp. 1–16, 2000.

[2] S. Barnard and M. Fisher, "Computational stereo," *ACM Comput. Surveys*, vol. 14, pp. 553–572, 1982.

[3] U. R. Dhond and J. K. Aggarwal, "Sructure from stereo - a review," *IEEE Trans. on Syst. Man and Cybern.*, vol. 19, pp. 1489–1510, 1989.

[4] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 993–1008, 2003.

[5] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo sorrespondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 7–42, 2002.

[6] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[7] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz, "Spacetime stereo: a unifying framework for depth from triangulation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 1–7, February 2005.

[8] L. Zhang, B. Curless, and S. M. Seitz, "Spacetime stereo: shape recovery for dynamic scenes," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Madison, WI, USA, 2003, pp. 367–374.

[9] M. Gong, "Enforcing temporal consistency in real-time stereo estimation," in *Proc. European Conference on Computer Vision*, Graz, Austria, May 7-13 2006, pp. III–564–577.

[10] M. El-Ansari, S. Mousset, and A. Bensrhair, "A new stereo matching approach for real-time road obstacle detection for situations with deteriorated visibility," in *Proc. IEEE Intelligent Vehicle Symposium*, Eindhoven University of Technology, Eindhoven, The Netherlands, June 4-6 2008.

[11] ——, "Temporal consistent real-time stereo for intelligent vehicles," *Pattern Recognition Letters*, vol. 31, no. 11, pp. 1226–1238, August 2010.

[12] M. El-Ansari, S. Mousset, A. Bensrhair, and G. Bebis, "Temporal consistent fast stereo matching for adavnced driver assistance systems (ADAS)," in *Proc. of IEEE Intelligent Vehciles Symposium*, San Diego, CA, USA, June 21-24 2010, pp. 825–831.

[13] Y. Otha and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, no. 2, pp. 139–154, 1989.

[14] H. Tao, H. S. Sawhney, and R. Kumar, "Dynamic depth recovery from multiple synchronized video streams," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Kauai, Hawai, USA, 2001.

[15] S. Vedula, S. Baker, P. Rander, R. Collins, and T. kanade, "Three-dimensioal scene flow," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, September 1991, pp. II–722–729.

[16] C. Leung, B. Appleton, B. C. Lovell, and C. Sun, "An energy minimisation approach to stereo-temporal dense reconstruction," in *Proc. of IEEE Int. Conf. on Pattern Recognition*, Cambridge, UK, 2004, pp. 72–75.

[17] G. Zhang, J. Jia, T. Wong, and H. Bao, "Consistent depth maps recovery from a video sequence," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 6, pp. 974–988, 2009.

[18] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, p. 679698, 1986.

[19] B. Cyganek and J. Borgosz, "An improved variogram analysis of the maximum expected disparity in stereo images," in *J. Bigun and T. Gustavsson (Eds.), SCIA 2003, LNCS 2749*, 2003, pp. 640–645.

[20] R. Labayrade, D. Aubert, and J. P. Tarel, "Real time obstacle detection in stereo vision on non flat road geometry through v-disparity representation," in *Proc. IEEE Intelligent Vehicle Symposium*, Versailles, France, June 2002.

[21] "Stereo data for algorithms evaluation," http://stereodatasets.wvandermak.com/, 2008.