

## **A disparity map refinement to enhance weakly-textured urban environment data**

Danilo Alves de Lima, Giovani Bernardes Vitor, Alessandro Corrêa Victorino,  
Janito Vaqueiro Ferreira

► **To cite this version:**

Danilo Alves de Lima, Giovani Bernardes Vitor, Alessandro Corrêa Victorino, Janito Vaqueiro Ferreira. A disparity map refinement to enhance weakly-textured urban environment data. IEEE International Conference on Advanced Robotics (ICAR 2013), Nov 2013, Montevideo, Uruguay. pp.1-6. hal-00936041

**HAL Id: hal-00936041**

**<https://hal.archives-ouvertes.fr/hal-00936041>**

Submitted on 30 Jan 2014

**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.

# A disparity map refinement to enhance weakly-textured urban environment data

Danilo A. Lima<sup>1</sup>, Giovani B. Vitor<sup>1,2</sup>, Alessandro C. Victorino<sup>1</sup> and Janito V. Ferreira<sup>2</sup>

**Abstract**—This paper presents an approach to refine noisy and sparse disparity maps from weakly-textured urban environments, enhancing their applicability in perception algorithms applied to autonomous vehicles urban navigation. Typically, the disparity maps are constructed by stereo matching techniques based on some image correlation algorithm. However, in urban environments with low texture variance elements, like asphalt pavements and shadows, the images' pixels are hard to match, which result in sparse and noisy disparity maps. In this work, the disparity map refinement will be performed by segmenting the reference image of the stereo system with a combination of filters and the Watershed transform to fit the formed clusters in planes with a RANSAC approach. The refined disparity map was processed with the KITTI flow benchmark achieving improvements in the final average error and data density.

**Index Terms**—Disparity map refinement, Computer Vision, Image Segmentation, Watershed Transform, RANSAC.

## I. INTRODUCTION

A car-like robot to perform safely its movement must be able to percept the neighbor environment, localize itself, plan its movement and control it. Each task has its specific problems, depending of the environment where the robot is inserted. Focusing on perception tasks for urban environments, the most common sensors are vision, LIDAR and sonar systems. Many of these sensors were used in the DARPA Grand Challenges, competitions held by the American's Defense Advanced Research Projects Agency (DARPA) between 2004 and 2007 to encourage the development of autonomous vehicles to perform tasks in a desert rally or in an urban environment. Several contributions have emerged from these DARPA's applications, like the advanced driver assistance systems (ADAS). Nowadays, these sensors still providing several improvements for the environment perception [1], where the number of sensors used is relevant to determine the viability for a real application.

As a low cost sensor, the stereo vision systems provide a large amount of data, depending of the camera field of view (FOV) and resolution. Several applications for urban environments perception use stereo cameras to detect free spaces and obstacles (Figure 1). However, they must deal with noisy data caused by the algorithms used to calculate the disparity map and the nature of environment features analyzed [2]. In urban environments, the most common noise sources are the weakly-textured elements as shadows, road texture, light variations, etc., which result in sparse



Fig. 1. TerraMax autonomous vehicle participant of the DARPA Urban Challenge which used a stereo vision system to urban environment perception [3].

disparity maps. Some works try to enhance the disparity map algorithms, proposing dense reconstructions, as the Global Matching techniques [4], producing good results with a high cost. As an intermediate solution are the Semi-Global Matching techniques [5], which present satisfactory results with fast embedded solutions [6] that allow applications for ADAS for example. Although these are robust techniques, the underestimation of the disparity map caused by weakly-textured elements occurs in some results and can compromise the accuracy of the detection of many free spaces and obstacles. To overcome these problems, some techniques apply the disparity map refinement to fit similar neighbor data in geometrical forms, like planes [7], [8]. The similar data are clusters grouped by local features (e.g., corners, edges, or colors) in the reference image.

This work address the problem of refining a Disparity map to improve its information in weakly-textured urban environments, where the result will be validated with the KITTI flow benchmark [9]. It uses a segmentation approach based on the Watershed Transform where the resulted clusters are fitted to planes, similar to the concept adopted in [10] to classify lines in the V-Disparity map. Differently from [11], the image segmentation with Watershed Transform was performed after preprocessing the reference image, providing a segmentation of the urban environment closer to its real elements boundaries, even in dark shadow areas. It is important to mention that there are several different applications for disparity maps in urban environments perception where the current refinement and segmentation approaches could be applied, as in road and obstacles detection [12], [13], [2], as well with techniques that incorporate more sensors to the system improving the vision capabilities to segment the traffic environments [14]. Another application in mind is the road profile extraction to aid the power consumption and dynamic model estimation of electric vehicles, as the one

The authors are with <sup>1</sup>Heudiasyc UMR CNRS 7253 Université de Technologie de Compiègne, <sup>2</sup>Universidade Estadual de Campinas (UNICAMP). Giovani B. Vitor holds a Ph.D. scholarship from CAPES/CNPQ and Danilo A. Lima holds a Ph.D scholarship from Picardi region. Contact authors danilo.alves-de-lima@hds.utc.fr

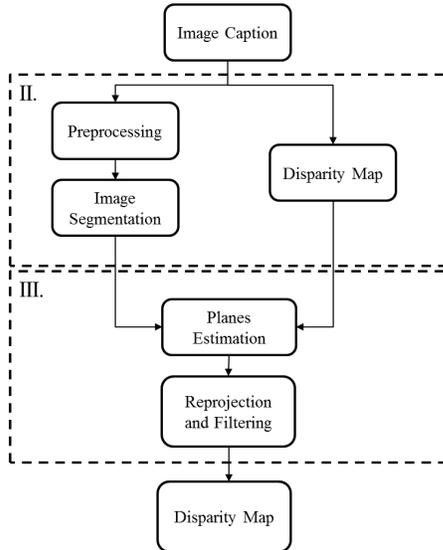


Fig. 2. Methodology block diagram.

proposed on the project VERVE<sup>1</sup>.

The Figure 2 presents a block diagram resuming the steps of our methodology. These blocks are detailed in the next sections of the paper, divided as follow: Section II presents the image segmentation process and the disparity map techniques used; Section III describes the planes estimation principle and the final point reprojection to compose the final disparity map; an experimental analysis and validation of the method, using sequences of the KITTI flow benchmark in an urban environment, is in Section IV; and, finally, Section V presents some conclusions and perspectives for future works.

## II. PREPROCESSING

The proposed methodology to refine the disparity map, Figure 2, starts with an image caption step. It uses a calibrated and synchronized stereo vision system which returns a stereo pair (two images) that are rectified to supply the following blocks, presented in the next subsections.

### A. Disparity Map

Following the diagram of the Figure 2, on the right are the Disparity Map block where the rectified images, based on the Epipolar geometry, are processed with some stereo matching algorithm to provide the disparity map ( $I_{\Delta}$ ) to be refined. To validate the presented solution, we used a local approach based on the Sum of Absolute Differences (SAD) correlation algorithm [15] and a Semi-Global Matching approach [5] to construct the  $I_{\Delta}$ . These current approaches are implemented in the OpenCV library as the Block Matching (BM) and the Semi-Global Block Matching (SGBM) algorithms and their results are already available in the KITTI stereo evaluation

<sup>1</sup>The project VERVE stands for *Novel Vehicle Dynamics Control Technique for Enhancing Active Safety and Range Extension of Intelligent Electric Vehicles*.

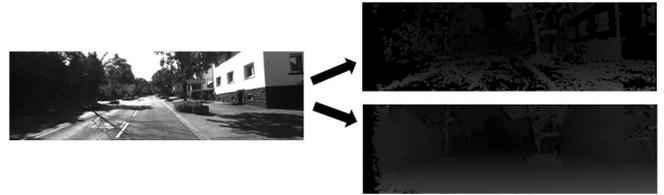


Fig. 3. Example of a rectified image of the left camera (reference image) of an urban view (left) and its respective disparity map with the BM (right top) and SGBM (right bottom) algorithms.

list<sup>2</sup>, justifying their use in this paper for evaluation purposes. They produce an sparse and dense  $I_{\Delta}$ , respectively, related to the percentage of valid points (non black pixels) as presented in the Figure 3. At the  $I_{\Delta}$ , the pixel intensity are the disparity values and they are related to the distance between the camera and the point in the world, given by  $Z = \frac{fB}{d}$ , where  $f$  is the focal length of the camera,  $B$  is the camera baseline, and  $d$  is the disparity value.

Several problems can be observed from these  $I_{\Delta}$  image, mainly caused by specific elements of the urban environments, such as shadows, light reflections, and low texture variances. These elements make difficult to find a right pixel matching in the stereo pair. Some of these matching problems are detected and eliminated as black points (0 value) in the  $I_{\Delta}$ . The remaining ones must be combined to estimate better values for them and the missing ones, as presented in the next steps.

### B. 2D Preprocessing

The conception of the Preprocessing block is to give more flexibility and support to the Image Segmentation. A image preprocessing is required when the segmentation works with the Watershed Transform, using local minima as marker [16]. The Watershed Transform is performed in a gradient image and its direct application produces a constraint of over-segmentation. Thus, three filters, the morphological gradient adjusted, the morphological reconstruction  $H_{min}$  and the morphological *Area Close*, are employed to avoid this constraint and providing the control of the segmentation level. This subsection explains these three algorithms used to preprocessing the reference image (Figure 3) before the segmentation.

The first one is the morphological gradient. Its formulation is given by the difference between dilation and erosion [17]:

$$gradMorph(f) = (f \oplus g_e) - (f \ominus g_i), \quad (1)$$

here  $f$  is the image function,  $g_e$  and  $g_i$  are structuring elements centered at the origin, and the operators  $\oplus$  and  $\ominus$  are respectively dilation and erosion (see [17] for more details).

In the early work [2], was observed that the low-contrast of higher frequency in shadow areas of the image provides

<sup>2</sup>Available in [http://www.cvlibs.net/datasets/kitti/eval\\_stereo\\_flow.php?benchmark=stereo](http://www.cvlibs.net/datasets/kitti/eval_stereo_flow.php?benchmark=stereo). Last access in 29 July 2013.

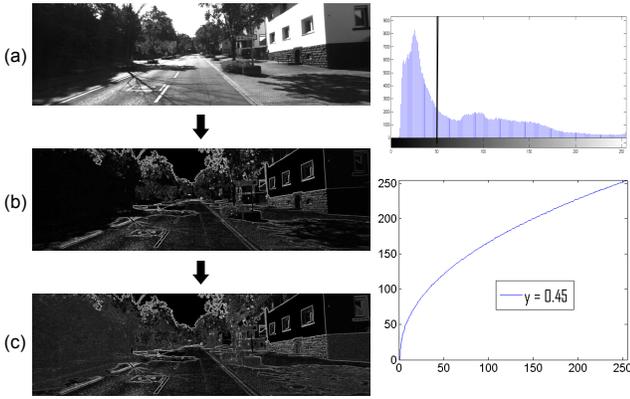


Fig. 4. Enhancing the contrast of higher frequency in shadow areas. (a) Original image, (b) gradient image and (c) the gradient image with shadow area enhanced.

a not correct segmentation by merging different regions. To avoid this drawback, it is proposed an enhancement on the areas where the shadow occur. It is performed applying a threshold in the grayscale original image to detect the shadow areas, where the grayscale value was set to 50 (to detect dark regions), and after is calculated a non-linear transform at the value of the gradient image increasing the contrast in this detected area. This transformation is given by the equation 2, which  $x$  represents the gradient intensity,  $c$  the constant of normalization and  $\gamma$  another constant factor defined to 0.45. These process is showed in the Figure 4.

$$f(x) = cx^\gamma. \quad (2)$$

After extracting higher frequencies from the reference image by the morphological gradient filter, the next step consists in using the *Area Close* to filter out pixels where their connected component area is smaller than a given parameter  $\lambda$ . Here is provided the definitions of *Area Open* once *Area Close* is its complement [18]. The connected component is defined as a pixels set characterized by a relation between their neighborhood on the binary image  $M \subset \mathbb{R}^2$ . First, the connected opening  $C_x(X)$  of a set  $X \subseteq M$  at point  $x \in X$  is the connected component of  $X$  containing  $x$  if  $x \in X$  and 0 otherwise. The binary area opening is then defined on subsets of  $M$  [19].

$$\gamma_\lambda^a(X) = \{x \in X \mid Area(C_x(X)) \geq \lambda\}. \quad (3)$$

The  $\gamma_\lambda^a(X)$  denote the morphological area opening with respect to the structure element  $a$  and the parameter  $\lambda$ . The  $Area(\cdot)$  is the number of elements in a connected component of  $C_x(X)$ . Its dual binary area closing is obtained as:

$$\phi_\lambda^a(X) = [\gamma_\lambda^a(X^c)]^c, \quad (4)$$

where  $X_c$  denotes the complement of  $X$  in  $M$ . Extending the filter for a mapping  $f : M \rightarrow \overline{\mathbb{R}}$ , in a grayscale image, then the area opening  $\gamma_\lambda^a(f)$  is given by:

$$(\gamma_\lambda^a(f))(x) = \sup\{h \leq f(x) \mid x \in \gamma_\lambda^a(T_h(f))\}. \quad (5)$$

In the Equation 5,  $T_h(f)$  represents the threshold of  $f$  at value  $h$ :

$$T_h(f) = \{x \in M \mid f(x) \geq h\}. \quad (6)$$

As mentioned before, the complement of equation 5 can be similarly extended to the conception of area closing to mappings from  $M \rightarrow \overline{\mathbb{R}}$ .

Finally, the last algorithm applied has the function of filter out the regional minima. It uses the morphological reconstruction  $H_{min}$ , obtained by successive geodesics dilations. This principle employs two subsets of  $\mathbb{R}^2$ , called *mask image* and *marker image*, with the same size. Moreover, the *mask image* must have intensity values higher than or equal to those from *marker image* [19]. Properly performing the reconstruction  $H$ -maxima or  $H_{max}$ , is possible to take the  $H$ -minima or  $H_{min}$  from its complement. Mathematically, defining *mask image* as  $I$  and *marker image* as  $I - h$ , being  $h$  the parameter to filter, the equation is given by [17]:

$$Hmax_h^a(I) = I \Delta_a (I - h). \quad (7)$$

In this definition,  $\Delta$  stands for morphological reconstruction with the structure element  $a$ . By duality, the  $H_{min}$  is defined as:

$$Hmin_h^a(I) = [I^c \Delta_a (I^c - h)]^c. \quad (8)$$

Studying different approaches, the Watershed Transform has different definitions at the literature, each one producing a different solution set, presented in [20]. The definitions are based on regional or global elements, such as influence zones and shortest-path forests with maximum or sum of weights of edges, or on local elements, such as the steepest descent paths. In this work, the local condition definition was used, called LC-WT (Local Condition Watershed Transform), which is defined as the steepest descent paths, where the neighbors information is used to create a path to the corresponding minimum. More details are found in [16].

The principle of this disparity map refinement is to associate the weakly-textured elements, like asphalt pavements and shadows, to segmented regions in the reference image, fitting a plane to each one of them, as well as, reduce the pixel variation in all regions. As mentioned, there are several works that use this analogy. In fact, what distinguishes this work is the way the image is segmented. The combination of these three filters with the Watershed Transform provides a different methodology to determine the resulted segmentation.

To demonstrate the processing done by this two blocks, the output of the segmentation can be seen in the Figure 5. Notice that the main point of this approach is choose the value of the parameters  $\lambda$  and  $h$  of the Preprocessing Block. In fact, these parameters give an excellent flexibility to determine the segmentation result of the Watershed Transform which is responsible to generate the representatives samples associated with the Disparity map.

### III. REMAPPING

#### A. Planes Estimation

Based on the assumption that pixels in a segment are coplanar, plane estimation is one of the most efficient ways for disparity refinement [8]. Similarly thought by [10] which found the navigable area by approximating small planes in the world by lines in the V-Disparity map, this work uses the RANSAC method to estimate plane coefficients on segments defined on the reference image with the points of the disparity map.

RANSAC based plane estimation is a minimization algorithm that can exclude the outliers where three points are randomly chosen to calculate the plane function during each iteration. Its accuracy is estimated by counting the number of points within the segment that agree with the given plane, like consensus. In this way, the plane fitting is performed with the points of the Disparity Map being  $P = (x, y, d)$  and  $C_p(s)$  being the set of segment points  $s$ . To each  $C_p(s)$ , the plane function estimated is given by the equation (9):

$$n \cdot (P - P_0) = 0, \quad (9)$$

where  $n$  is the nonzero vector normal to the plane,  $P_0$  is a know point in the plane,  $P$  is a given point to determine if it is in the plane, and  $(\cdot)$  means a dot product.

#### B. Disparity Map Reprojection and Filtering

For each cluster's fitted plane, the Disparity image ( $I_\Delta$ ) can be reprojected applying its points in the respective plane equation. However, due to the nature of the RANSAC and the nonlinear variation of the disparity data, the final plane is only a estimation that can be good or not, depending of the points used in the process. Therefore, the  $I_\Delta$ 's points must respect the follow statements to be fit:

- The pixel disparity value is in a interval where a small variation in its value does not represents a large change in the Z distance, in the world (e.g., variation bigger than 2 meters, based in the relation presented in II-A);
- The segment have enough percentage of valid points (disparity value different of zero);
- The maximum number of interactions to fit the plane was not reached;
- The point reprojected do not highly diverge from its original value (which is different from zero).

The highly divergence mentioned in the last statement is determined by the difference between the disparity value in  $I_\Delta$  and the reprojected disparity value. Following these statements the Refined Disparity map  $I_{\Delta_r}$  is acquired.

### IV. EXPERIMENTAL RESULTS

In this section is presented the results for the Disparity map refinement, demonstrating the influence of the parameters  $\lambda$  and  $H$  on the image segmentation resulted, as well as its validation in the final result of the proposed method. The methodology was tested with the KITTI flow benchmark [9] acquired under a static environment, including 194 training and 195 test image pairs of diversified urban

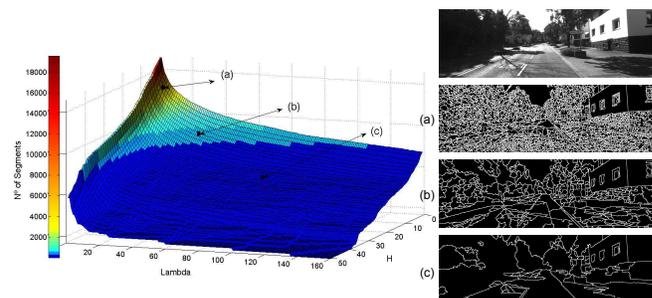


Fig. 5. Example of an influence surface for the parameters  $\lambda$  and  $h$  on the number of segments in a given image. In (a)  $\lambda_1 = 5$ ,  $h_1 = 2$  and 4090 segments; (b)  $\lambda_2 = 30$ ,  $h_2 = 5$  and 427 segments; (c)  $\lambda_3 = 80$ ,  $h_3 = 15$  and 73 segments.

scenarios, with a ground truth with approximately 30% of pixels coverage in the disparity map. The complete loop without any code optimization is around 1.7s in an Intel CPU I7 with 2.6GHz clock and 8GB of RAM at the current stage. After validated the proposed solution, the performance can be better optimized by software and hardware solutions, as GPU cards [21].

The Figure 5 demonstrates the number of segments resulted in a given image, ranging the  $\lambda$  parameter from 0 to 160 units and the  $H$  parameter from 0 to 60 units. It can be seen in the graphic that the number of segments fall exponentially in the sense of increase the two parameters. In the Figure 5 can also be seen the result of three combination of  $\lambda$  and  $h$  values, where in (a)  $\lambda_1 = 5$  and  $h_1 = 2$ , obtaining 4090 segments, in (b) was applied  $\lambda_2 = 30$  and  $h_2 = 5$ , obtaining 427 segments, and the last one in (c) setting  $\lambda_3 = 80$  and  $h_3 = 15$ , obtaining 73 segments.

To evaluate the influence of these  $\lambda$ 's and  $h$ 's values in the refinement of the Disparity Map  $I_\Delta$  calculated by the Block Matching (BM) and the Semi-Global Block Matching (SGBM) algorithms ([15], [5]), the methodology presented in the Figure 2 was executed for each training image. The KITTI benchmark provides an evaluation algorithm that returns three metrics: the average number of pixels that diverges of the ground truth (non occluded and all), the end-point error that represents the mean variation for each pixel, and the density of valid points. The original  $I_\Delta$ 's were acquired with the same configuration used in the KITTI stereo evaluation list, where the BM used has the 28<sup>th</sup> position and the SGBM has the 20<sup>th</sup> position. The complete result is in the table I. In this table, the highlighted rows present the best combinations of  $\lambda$  and  $H$  which improve the original  $I_\Delta$  method. It is possible to see meaningful reductions in the final pixel error with expressive gains in the density of valid points. In the BM refinement case, where the original  $I_\Delta$ 's give less than 50 % of valid points to planes estimation, this estimation will be more sensible to variations in these points. Nevertheless, the final  $I_{\Delta_r}$ 's average pixels variation error are not increased as much as they improved their density. However, in the case of the SGBM refinement, where more pixels are available to estimate the plane, the final  $I_{\Delta_r}$ 's reduce all the parameters analyzed.

TABLE I  
FINAL REFINEMENT RESULT FOR THE KITTI TRAINING DATASET [9]

Method	> 2 pixels		> 3 pixels		> 4 pixels		End-Point		Density
	Out-Noc	Out-All	Out-Noc	Out-All	Out-Noc	Out-All	Out-Noc	Out-All	
BM	17.09 %	18.34 %	12.54 %	13.68 %	10.15 %	11.20 %	2.26 px	2.58 px	47.48 %
BM + Our( $\lambda_1, H_1$ )	19.21 %	20.45 %	13.80 %	14.97 %	10.80 %	11.88 %	2.27 px	2.61 px	68.29 %
<b>BM + Our(<math>\lambda_2, H_2</math>)</b>	<b>17.78 %</b>	<b>18.97 %</b>	<b>13.28 %</b>	<b>14.38 %</b>	<b>10.68 %</b>	<b>11.70 %</b>	<b>2.19 px</b>	<b>2.52 px</b>	<b>77.24 %</b>
BM + Our( $\lambda_3, H_3$ )	24.61 %	25.59 %	19.01 %	19.93 %	15.58 %	16.44 %	2.84 px	3.10 px	85.82 %
SGBM	16.86 %	18.51 %	13.52 %	15.12 %	11.69 %	13.25 %	3.15 px	3.87 px	87.31 %
<b>SGBM + Our(<math>\lambda_1, H_1</math>)</b>	<b>15.72 %</b>	<b>17.40 %</b>	<b>12.01 %</b>	<b>13.68 %</b>	<b>10.02 %</b>	<b>11.65 %</b>	<b>2.59 px</b>	<b>3.32 px</b>	<b>95.90 %</b>
<b>SGBM + Our(<math>\lambda_2, H_2</math>)</b>	<b>16.18 %</b>	<b>17.60 %</b>	<b>12.28 %</b>	<b>13.66 %</b>	<b>10.04 %</b>	<b>11.37 %</b>	<b>2.31 px</b>	<b>2.88 px</b>	<b>97.80 %</b>
SGBM + Our( $\lambda_3, H_3$ )	22.12 %	23.25 %	16.84 %	17.94 %	13.52 %	14.60 %	2.68 px	3.13 px	96.68 %

The Figure 6 presents general results for the proposed disparity map refinement methodology with the BM approach, comparing it to the original disparity map. It is possible to see how the refinement enhance the disparity information (density of valid points) and the final error is reduced in regions with well defined pixels. The road surface, for example, with many invalid points, had a better estimation by this approach. Similar description was proposed for the SGBM approach, resulting the Figure 7.

## V. CONCLUSIONS AND FUTURE WORKS

This work presented a disparity map refinement applied to weakly-textured urban environments. The solution used a noisy and sparse disparity map information to fit planes in segments defined in the reference image, a similar concept used in the v-disparity map to detect lines and approximate the world to small planes. These segments were created by a 2D image segmentation based on the combination of three filters and the Watershed Transform. The final result was tested with the KITTI benchmark, providing a solid validation for the present approach.

The proposed refinement improved the original disparity map fitting the missing and noisy information, which belong to the same defined segment and agree with the conditions of a reliable qualitative result observed in the KITTI benchmark. This new information added also improves the visualization of small size elements, such as the difference between the road and the side-walk, once the pixel variation error was reduced. However, there are situations where the segments did not have enough data to estimate a plane, keeping higher pixel errors. These differences in the plane estimation were evidenced in the BM and SGBM disparity maps refinement.

To take advantage of this disparity map refinement, different approaches will be analyzed as an improvement for world semantic labeling, with obstacles detection and segmentation, as well as curbs detection. The final intention is to use this entire information as a perception layer of the autonomous vehicle from the project VERVE. To acquire this, the processing time of each step of the Figure 2 will be worked, using GPU cards applications, such as the one presented in [21].

## REFERENCES

- [1] J. R. McBride, J. C. Ivan, D. S. Rhode, J. D. Rupp, M. Y. Rupp, J. D. Higgins, D. D. Turner, and R. M. Eustice, "A perspective on emerging automotive safety applications, derived from lessons learned through participation in the darpa grand challenges," *Journal of Field Robotics*, vol. 25, pp. 808–840, October 2008.
- [2] G. B. Vitor, D. A. Lima, A. C. Victorino, and J. V. Ferreira, "A 2d/3d vision based approach applied to road detection in urban environments," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013, pp. 952–957.
- [3] Y.-L. Chen, V. Sundareswaran, C. Anderson, A. Broggi, P. Grisleri, P. Porta, P. Zani, and J. Beck, "Terramax: Team oshkosh urban robot," in *The DARPA Urban Challenge*, ser. Springer Tracts in Advanced Robotics, M. Buehler, K. Iagnemma, and S. Singh, Eds. Springer Berlin Heidelberg, 2009, vol. 56, pp. 595–622. [Online]. Available: [http://dx.doi.org/10.1007/978-3-642-03991-1\\_14](http://dx.doi.org/10.1007/978-3-642-03991-1_14)
- [4] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 2, 2001, pp. 508–515 vol.2.
- [5] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, 2005, pp. 807–814 vol. 2.
- [6] C. Banz, S. Hesselbarth, H. Flatt, H. Blume, and P. Pirsch, "Real-time stereo vision system using semi-global matching disparity estimation: Architecture and fpga-implementation," in *Embedded Computer Systems (SAMOS), 2010 International Conference on*, 2010, pp. 93–101.
- [7] R. Gupta and S. Y. Cho, "A color-based approach for disparity refinement," in *Control Automation Robotics Vision (ICARCV), 2010 11th International Conference on*, dec. 2010, pp. 664–667.
- [8] L. Xu, O. Au, W. Sun, Y. Li, and J. Li, "Hybrid plane fitting for depth estimation," in *Signal Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, dec. 2012, pp. 1–4.
- [9] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 3354–3361.
- [10] R. Labayrade, D. Aubert, and J. P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," in *Proceedings of the IEEE Symposium on Intelligent Vehicles*, vol. 2, 2002, pp. 646–651.
- [11] K. Hajebi and J. Zelek, "Structure from infrared stereo images," in *Computer and Robot Vision, 2008. CRV '08. Canadian Conference on*, may 2008, pp. 105–112.
- [12] N. Soquet, D. Aubert, and N. Hautiere, "Road segmentation supervised by an extended v-disparity algorithm for autonomous navigation," in *Proceedings of the IEEE Symposium on Intelligent Vehicles*, 2007, pp. 160–165.
- [13] F. Dornaika, J. Alvarez, A. Sappa, and A. Lopez, "A new framework for stereo sensor pose through road segmentation and registration," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 4, pp. 954–966, dec. 2011.
- [14] M. Teutsch, T. Heger, T. Schamm, and J. Zollner, "3d-segmentation of traffic environments with u/v-disparity supported by radar-given



Fig. 6. Result for the Disparity map refinement of the BM method in different urban conditions.

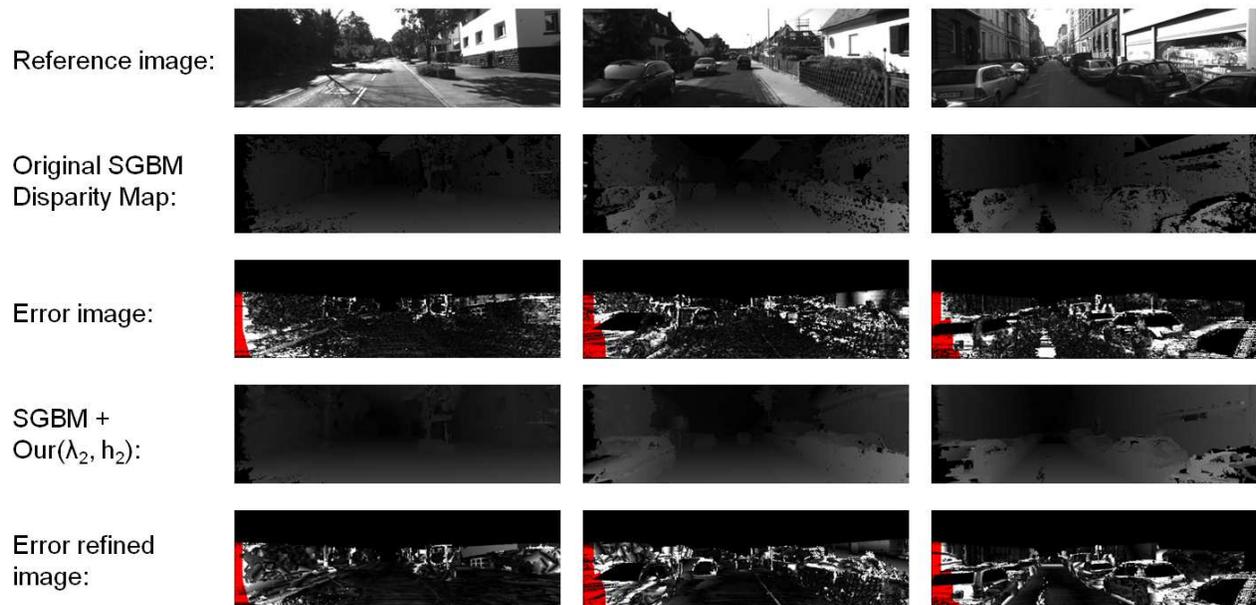


Fig. 7. Result for the Disparity map refinement of the SGBM method in different urban conditions.

- masterpoints," in *Intelligent Vehicles Symposium (IV)*, 2010 IEEE, June 2010, pp. 787–792.
- [15] T. K. Hiroshi, H. Kano, S. Kimura, A. Yoshida, and K. Oda, "Development of a video-rate stereo machine," in *IROS '95: Proceedings of the International Conference on Intelligent Robots and Systems-Volume 3*. Washington, DC, USA: IEEE Computer Society, 1995, p. 3095.
- [16] R. Audigier and R. de Alencar Lotufo, "Relationships between some watershed definitions and their tie-zone transforms," *Image Vision Comput.*, vol. 28, no. 10, pp. 1472–1482, Oct. 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.imavis.2009.11.002>
- [17] E. R. Dougherty and R. A. Lotufo, *Hands-on Morphological Image Processing (SPIE Tutorial Texts in Optical Engineering Vol. TT59)*. SPIE Publications, Jul. 2003.
- [18] A. Meijster and M. H. F. Wilkinson, "A comparison of algorithms for connected set openings and closings," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 484–494, 2002.
- [19] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *Image Processing, IEEE Transactions on*, vol. 2, no. 2, pp. 176–201, Apr.
- [20] J. B. T. M. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," 2001.
- [21] A. Korbes, G. Vitor, R. de Alencar Lotufo, and J. Ferreira, "Advances on watershed processing on gpu architecture," in *Mathematical Morphology and Its Applications to Image and Signal Processing*, ser. Lecture Notes in Computer Science, P. Soille, M. Pesaresi, and G. Ouzounis, Eds. Springer Berlin / Heidelberg, 2011, vol. 6671, pp. 260–271.