Disparity Map Adjustment: a Post-Processing Technique

Gabriel da Silva Vieira[1*§], Fabrizzio Alphonsus A.M.N. Soares[2*], Gustavo T. Laureano[3*], Rafael T. Parreira[4*], Júlio C. Ferreira[5§], Rogério Salvini.[6*]
[*]Federal University of Goiás, *Pixelab Laboratory*. Goiânia - GO, Brazil.
{fabrizzio[2], gustavo[3], rogeriosalvini[6]}@inf.ufg.br, rafaeltp3[4]@gmail.com
[§]Federal Institute Goiano, *Computer Vision Laboratory*. Urutaí - GO, Brazil.
{gabriel.vieira[1], julio.ferreira[5]}@ifgoiano.edu.br

***Abstract*—As a digital image provides such information about a scene, a disparity map can be yielded by means of stereo images. This topic was exhaustively surveyed but it remains one of the most important branches in both computer vision and machine vision. Most algorithms are organized in a pipeline that starts with a matching cost step and ends with a disparity refinement. This paper provides a simple but an effective method to adjust a disparity map in a more appropriate configuration, i.e. it presents a disparity refinement technique. It is based on an assumption that most disparities in a region point to a correct disparity value for this area. To develop the methodology, we use image segmentation and support weighted windows. By performing an evaluation, it shows that this method can increase the robustness of a raw disparity map even with a lot of noisy parts.**

Keywords: stereo vision; adaptive support window; disparity map; map refinement; disparity methods; image segmentation.

I. Introduction

Depth recovery is a primary problem in several fields. For instance, an autonomous vehicle can use the depth of a scene to move around without collisions. A visually impaired person can use a computer application developed to help him in his mobility. Objects can be tracked with no overlap. Augmented reality can bring better visual effects and so on.

As a digital image provides such information about a scene, it can also be used for estimating the position of objects in space, which can be accomplished by stereo imaging. Thus, for obtaining a depth map, digital images are used to calculate the disparity between similar points.

A disparity map can be yielded by means of stereo images. When two or more cameras are used to capture the same scene, the displacement of the cameras produce a paralaxe which puts a point in different coordinates of each image plane. The movement of a point can be followed and its motion shows how far away a point has changed its position. Hence, when this displacement is calculated we obtain the disparity among correspondent points.

In case that all points in a scene have been followed, a dense disparity map can be produced. On the other hand, if just some points have been considered a sparse disparity map is returned. Depending on the application, a sparse or dense map can be chosen. For instance, for some types of autonomous robot just points with high disparities that are important. In this case, a sparse map can be considered whereas points with low disparities may be discarded. However, an application which aims to mapping in details a 3D scene, presumably, it will use a dense map.

Disparity maps are used to estimate the depth relation among objects in a scene and the depth relation between them and the camera equipment. Thus, stereo vision systems aim at inferring disparity maps by processing a pair of images, in a binocular approach, or by considering three or more images, in a multiview approach.

Accordingly to Mattocia et al. [1], this topic was exhaustively surveyed. Despite that fact, in a more recent work, Hamzah and Ibrahim [2] show that the field of stereo vision remains active in research and development. Yang et al. [3] endorse it by affirming that nowadays stereo vision is one of the most important branches in both computer vision and machine vision with application in different types of fields, as 3D shap measurement and target tracking.

In stereo vision systems, most algorithms are organized in four steps: (1) *matching cost*, (2) *cost aggregation*, (3) *disparity selection* and (4) *disparity refinement*. This pipeline was pointed out by [4] and a lot of work has been done based on it.

At the matching cost step, a mathematical model is defined to measure the dissimilarity value between points. Models like that are known in the literature as cost functions, match measures or even as photo-consistency. Absolute Differences (AD), Squared Differences (SD) and Truncated Absolute Differences (TAD) are often reminded.

The cost aggregation step is mainly used in local stereo methodologies, in opposition of global strategies. In local methods, the matching cost is calculated over a neighborhood region which is commonly referred to as support or aggregating window. They consider the entire set of pixels associated with the image regions that may be square or rectangular and may be fixed or adaptive in size [2].

The disparity selection step deals with the task of choose the most suitable disparity for each point, a common strategy is the Winner Takes All (WTA). Finally, at the disparity refinement step, a post-processing technique is applied in the disparity map to adjust disparities which were wrongly estimated.

Our study focuses on step 4 of the pipeline, which is the *disparity refinement*. At this work, we propose a simple but an effective technique to adjust a disparity map in a more appropriate configuration. It is based on an assumption that disparity map regions carry helpful information. As a set of disparity are allocated in a region, the disparity with more frequency, i.e, the disparity that mostly appears in a region, indicates the correct one, with a condition that a region has similar points. Hence, other disparities in that region can be discarded. The next step is to find the disparity of these points. To do that a weighted function is performed for this task.

By performing an evaluation, we can observe that this strategy can improve disparity maps substantially. Even maps with low accuracy can be enhanced by this method. In addition, we pointed out that although this proposal is presented in the context of local approaches, it can also be applied in global strategies in the same way.

The remainder of the paper is organized as follows. After briefly reviewing closely related work in Section II, we show a overview of the technique proposed and we describe our algorithm for adjusting a disparity map in Section III. The experimental results and analysis are given in Section IV and Section V concludes the paper.

II. Related Work

When we observe the field of research in stereo vision systems, the most prominent area in the disparity map building is the matching cost step (number 2 in the pipeline). Different methodologies have been proposed, each one with a new idea or as an incremental approach. Such studies have made a solid base that supports computer vision applications in a variety of real world problems.

Although the number of researches in stereo vision are amazingly high, keeping up with the new trends is the hardest part. Hamzah and Ibrahim [2], in a recent survey, showed that almost 200 original papers are published per year only in IEEE Xplore database. This fact was also observed in an older survey. Lazaros et al. [5], when analysing the stereo vision field, pointed out that new approaches were being presented every year and such expanding volume of work was making it difficult for those interested to keep up with it.

Kumaru and Kaur [6], by discussing this point, concluded that though a large number of methods have been developed for calculating disparity, the problem is still ill-posed with major challenges as photometric variations, untextured/repetitive regions and high density of noise.

These problems are also common challenges in the other disparity map pipeline steps. If a difficult task is prolonged, then it may be treated in another moment. Thus, at some point, this protracted task will culminate in the refinement step. Whereas the final goal is to obtain an accurate map, this last step is often applied to obtain robustness.

The refinement procedure takes a raw disparity map to perform some manipulation with this data. Thus, a data analysis is conducted to define a methodology capable to deal with some considered aspects, as border preserving and occlusion handling.

Depending on the quality of this raw map, a final result may achieve considerable success. It is because an adjustment takes this raw map into account, so to produce a robust map from a degraded one is not a trivial question.

This adjustment envolves a set of procedures which are based on hypotheses, heuristics or simple observations. However, unlike the matching step, the refinement one is not very often discussed, at least with emphasis. Proposals are made but they are presented briefly in a section of a paper, i.e, they are put in the background.

Left to right consistency-check is a frequent refinement process that is applied in a raw disparity map. It consists of cross checking two or more disparity maps. For instance, to apply this methodology in a binocular approach, two disparity maps must be yielded, so each one of the image pair is used as a reference image, one at a time.

This consistency-check method deals with areas that are occluded. Points that are not visible may be detected and labeled as unknown disparities. After that, another technique can be used to fill in these occluded pixels. This method can improve a raw map by observing neighboring points. Therefore, if a disparity of a point is unknown, probably the neighboring points may tell what is.

Yang et al. [3] explains this method. The disparities of occluded pixels on the left disparity map are assigned to zeros. The occluded pixels on the left disparity map are filled with the lowest ones of their horizontal adjacent disparities or with a weighted median filter that is further used to eliminate the stretching effect caused in the horizontal filling.

The same is explained by Rhemann et al. [7]. When an occluded pixel is detected, it is assigned to the lowest disparity value of the spatially closest non-occluded pixels which lie on the same scanline (pixel row). However, they point out that this simple occlusion filling strategy can generate streak-like artifacts in the disparity map. Thus to remove them, while preserving the object boundaries, a weighted median filter can be applied to the filled pixels.

This method is not a recent one. In a classic stereo vision paper, Scharstein and Szeliski [4] presented it by saying that occluded areas can be detected using cross-checking (comparing left-to-right and right-to-left disparity maps). Although they didn't use it in their experiments, they confirm this method can be applied to remove spurious mismatches.

Hosni et al. [8] used this method with a simple modification. They observed that since occlusion occurs in the background of an image, the occluded pixel can be assigned to the minimum value from both disparity maps. According to them, this strategy also generates horizontal streaks in the disparity map and hence it demands post-processing on the filled in pixels.

This disparity map combination was also used by [9]. To improve the accuracy of their results, they calculated a depth image for both stereo images and combined them to eliminate some final artifacts. They assumed that artifact faults only occur in one of the two views, so they took the minimum of both disparity maps.

According to Hirschmuller and Scharstein [10], the refinement step is applied to reduce the overall errors, which in turn yields improvements of the final result. They also performed the left to right consistency-check in their experiments. It was used for invalidating occlusions and mismatches where invalid disparity areas were filled by propagating neighboring small disparity values.

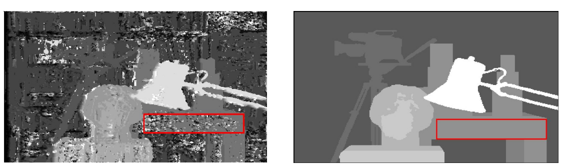
Apart from that method, other methodologies can be found. Hirschmuller et al. [11] proposed a border correction filter that modifies the disparity image by horizontally shifting assumed object borders. Wang et al. [12] developed an algorithm using the MRF framework, image inpainting technique and image color contrasts to eliminate holes and misaligned pixels. Hirschmuller [13] presented post-processing steps for removing wrong disparities, recovering from specific problems of structured environments, and the interpolation of gaps for preserving discontinuity.

Mattocia el at. [14] proposed a powerful method that uses adaptive weights for classifying pixels based on geometric and photometric constraints. It takes a pair of images and a raw disparity map as input, then the plausibility of each point is evaluated by considering the relation among points in the same aggregating window, points between images and the original disparity.

In this brief section, some strategies of disparity enhancement were pointed out. It doesn't exclude other methodologies but it shows that the most discussed method in this literature is the left to right consistency-check. This is most likely due to its simplicity for implementing and ability to find occluded points with efficiency. However, other methods are proposed but sometimes with no great significance.

III. Proposed Approach

We start by analysing a raw disparity map. Fig. 1 shows a map that is very noisy in some parts of it. It was made by a simple cost aggregating (CA) methodology that can be called as *fixed window* (FW) method. It is the simplest CA strategy that uses an aggregating window and it is at the foundation of stereo vision systems. Besides, this map was also yielded by using a simple cost function that is the *sum of absolute differences* (SAD).



(a) (b)

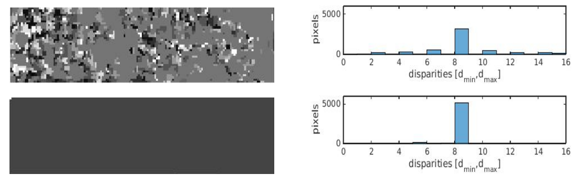
Fig. 1: Disparity maps: (a) a raw map and (b) a ground truth map.

This simple form, made by FW method and SAD cost, structures a method that has some advantages. First, it is easy to implement. Second, it runs fast in a computer. Third, it is parallelizable. Four, by using a small window it preserves borders and by using large windows it reduces noise.

On the other hand it has some disadvantages. One of them is that it fails in textureless regions. It consideres that points at the same window have the same disparity. Furthermore, the best window size needs to be found empirically and it produces noisy maps caused by mismatches.

In spite of the disadvantages, we take the advantages as a motivation. Thus, our strategy is prepared by considering a raw disparity map which has noisy parts. Probably, a better disparity map, yielded by a robust stereo method, produces better results but the simplicity of the FW method and its advantages justify an investigation.

Fig. 1a illustrates a region that has a group of wrong disparities. Similar pixels coexist in this area and because of that, FW method fails in a lot of points. However, when we analyse these disparities we can see that most of the values are pointing to a correct one. Fig. 2 shows an histogram plot which confirms our analyse by comparing this map region with the same region in the reference map (*ground truth*), Fig. 1b.

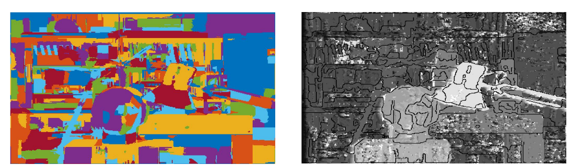


(a) (b)

Fig. 2: Map analysis: (a) disparity map and (b) related histograms.

In this way, if in a certain region a disparity method hits more than fails, we can use it. Unfortunately, it is something that we don't know because the correct disparity is still unknown. But if we believe in it, we can propagate this supposed correct value even knowing that this is not true all the time. Our methodology starts with this belief.

To identify a region, a segmentation technique may be used. In stereo vision systems, mean shift algorithm [15] is widely employed. It was used to obtain great results in [9], [16] and [17]. We use it to apply a segmentation in the reference image. When we obtain these segments we use them to localize regions in the disparity map. Fig. 3 shows a segmented image and its corresponding disparity map labeled based on these segments.



(a) (b)

Fig. 3: Image segmentation: (a) reference image and (b) disparity map.

After that, the method calculates the most common value for each segment. It is a simple equation that is show in Eq. 1. For each segment S with the identifier i into the disparity map D, $(S_i \subset D)$, it calculates the mode of all n segments and the results are stored in m.

$$m_i = mode\ common\ value\ in\ S_{i=1}^{n} \qquad (1)$$

Moreover, each point of the disparity map that belongs a certain segment is evaluated, accordingly with the previous mode contability. In Eq. 2, a disparity value in D with the coordinates (x, y) is tested. In case of this value is in a range test, the mode value m is assigned for this point. Otherwise, it is assigned with 0 that represents a unknown disparity. In this equation, t is a threshould defined by a user that is used to approximate disparity values to the segment's mode. Besides, it considers that each disparity value is in a segment S with the identifier i.

$$D(x,y) = \begin{cases} m_i & \text{if } D(x,y)_{\in\{S_i\}} \in [m_i - t, m_i + t], \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

When applying the above equations, a disparity map is returned. At this time, disparities that are far away from their segment mode value are considered as unknown. The next step consists of filling these holes in, so a weighted filter is prepared to evaluate the plausibility of each possible disparity.

Yoon and Kweon [18] introduced a support weighted window to be applied in the stereo matching problem. Their methodology considers the color similarity between points and their space distance. A window is defined and a point located in the middle of this window is the principal point. The surrounding neighbors are compared with the principal point by calculating their difference of colors and their geometric distance. This strategy was used in [14], [19], [20] among others and investigated in [8].

The color proximity constraint between a principal point p and its neighbor point n within a support is given by:

$$f_c(\Delta c_{pn}) = e^{-\frac{\Delta c_{pn}}{\gamma c}} \quad (3)$$

The color distance Δc_{pn} represents the Euclidean distance between the colors of p and n in an image I as

$$\Delta c_{pn} = \sqrt{\sum_{j \in r,g,b} (I_j(p) - I_j(n))^2} \quad (4)$$

In the same way, spatial proximity constraint is evaluated accordingly to:

$$f_s(\Delta s_{pn}) = e^{-\frac{\Delta s_{pn}}{\gamma s}} \quad (5)$$

the spatial distance Δs_{pn} represents the Euclidean distance between the coordinates (x, y) of p and n as

$$\Delta s_{pn} = \sqrt{(p_x - n_x)^2 + (p_y - n_y)^2} \quad (6)$$

γc and γs refer to a constant of color similarity and a constant to adjust the spatial distance term, respectively. $f_c(\Delta c_{pn})$ and $f_s(\Delta s_{pn})$ represent the strength of grouping by color similarity and by proximity.

Color and spatial constraints are combined and the final support weighted window is given by

$$W(p,n) = e^{-(\frac{\Delta c_{pn}}{\gamma c} + \frac{\Delta s_{pn}}{\gamma s})} \quad (7)$$

Hosni et al. [8] investigated the Eq. 7 and showed that the spatial constraint can be omitted with very little difference on the quality of results, as

$$W(p,n) = e^{-(\frac{\Delta c_{pn}}{\gamma c})} \quad (8)$$

In our method, we use the support weighted window with an adaptation. It is only applied in unknown disparities so a principal point in a window is a point of disparity that we want to discover. Each neighboring pixel that has a disparity value is evaluated according to the previous equations. Thus, the weights of each pixel that are in the same disparity are accumulated. Fig. 4 helps in the explanation.

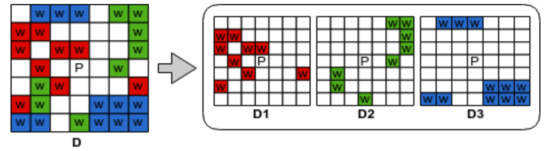


Fig. 4: Points in a disparity map D have their own weight w. Each segment is painted didactically (in red, green and blue). Weights in $D1$, $D2$ and $D3$ are summed separately. The best value is used to set the disparity in the principal point p.

Based on the color reference image, the photometric and geometric constraints are calculated and each point of the window has a weight w. Besides the weights, we know some disparities. In Fig. 4, each color represents a known disparity, except for a white color point that represents an unknown disparity and because of that these points don't have a weight w. Thus, the computed weights that are in the same disparity are summed up as in

$$\Omega_{d \in \{d_{min}, d_{max}\}} = \sum_{j=1}^{n} w_{ij} \quad (9)$$

where d_{min} and d_{max} are the range from minimum to maximum disparity and Ω is the accumulated sums. Hence, a disparity optimization is performed to select the best disparity. It is given by

$$D_{(x,y)} = \text{argmax}(\Omega) \quad (10)$$

where (x, y) are the coordinates of the unknown disparity in the disparity map D. Based on the best value from Ω, its disparity value is assigned to D.

IV. Experimental Results

In this section, the results and the organization of the experiment are presented. Four image pairs were selected from the Middlebury dataset [21]. Each pair has its own ground truth that was used to evaluate the results. The methodology followed the specifications of [4].

In Table I, parameters ALL, NOCC and DISC are defined according to the Middlebury Stereo Evaluation - version 2 [21]. Although this version is no longer active, it is still being used, as in [22]. Some evaluation masks are provided and they are used to remove pixels that are not considered in the statistics. ALL is the error computed on the whole image, NOCC is the error computed on the whole image excluding the occluded regions and DISC is the error computed within the discontinuity regions [14].

In the test cases, raw disparity maps from FW method are used as input for the proposed methodology, referred to as *segment consistency-check* (SCC).

A 3×3 aggregation window was used to build raw disparity maps and the SAD cost was used as a measure of stereo matching. The SCC method used $t = 1$ for Eq. 2, $\gamma c = 23$ and $\gamma s = 14$ for Eq. 7 (and for Eq. 8).

In our experiment, we test both strategies to fill in unknown disparities: with two constraints (color and spatial proximity) and with only a constraint (color and no spatial proximity). In case of using no spatial constraint the method is referred as SCCNoSpatial.

Besides that, a 39×39 window was defined for the support weighted function. To find the best window size and its corresponding γs term, we followed the methodology proposed by [20]. The best parameters found for Tsukuba image pair were used in all tests.

Table I shows the accuracy of the proposed method. The SCC method decreased the percentage of bad pixels in the three considered parameters. For instance, the bad pixel error in Tsukuba image pair was reduced substantially from 20.81 to 4.41 percent in the ALL parameter. The same occured in NOCC and DISC parameters.

Table I still shows that SCCNoSpatial obtained important results, especially in Tsukuba and Cones images. However SCC with both constraints obtained better results in DISC parameter.

The experiments were made in a notebook with Core i5-460M 2.53 GHz CPU and 4 GB RAM, and no parallelism technique was utilized. As for the execution time, the whole process (FW + SCC) required 27.81 seconds on Tsukuba (0.53 and 27.28 seconds, respectively). In the other image pairs, we report 51.63 seconds for Venus, 75.05 for Teddy and 75.08 for Cones in the overall process.

A qualitative analysis shows an evident improvement brought in by the proposed methodology compared to the raw disparity maps. Fig. 5 shows the disparity maps to each image pair and the disparity map improvement with the SCC method. The first column from Figure 5 corresponds to the original image. The raw disparity maps are in the second column. In the third column, there are maps after applying Eq. 2. The fourth column shows the final result and the last column displays the ground truth.

We extended this evaluation and it was applied in a disparity map produced in a pixel based technique (i.e, without the second step in the stereo vision pipeline). Table II shows the results. Again, SCC method performed well and increased the robustness by reducing the disparity error. A visual result is shown in Fig. 6.

V. Conclusion and Future Work

The stereo vision field is an interesting and challenging area of research. Surveys have shown that a considerable number of researches are being developed and a set of papers are being produced every year.

One of them was prepared by Scharstein and Szeliski [4] which brought a taxonomy, evaluation metrics, rectified stereo images and ground truth images within a web platform. The pipeline proposed by the authors has influenced a lot of work especially in the cost aggregation step.

However, in this work we explore step 4 which is the disparity refinement. A methodology was proposed and experiments were performed. Our approach is based on disparity points within a segment that can reveal the appropriate disparity value for this area. Thus, the most common disparities can be propagated. By doing this we eliminate possible wrong disparities and we used a support weighted window to find the best disparity for the unknowns.

Experimental results show that this method can increase the robustness of a raw disparity map even with a lot of noisy parts. Thus, when the SCC method is applied in a raw map produced by the *fixed window* (FW) method, the disparity errors are reduced in an important way as to be worthy of attention, as shown in Tables I and II.

In the next phase, we intend to perform the SCC method in a more recent dataset as version 3 of the Middlebury Stereo Evaluation which has more complex scenes. Besides, raw disparity maps produced by different methods can be investigated, for instance, maps produced by adaptive support windows methods. Thus, we could observe whether the SCC method has major influence in their results. Furthermore, inpainting techniques could be investigated to select the best disparity in big unknown areas. Likewise, textureless regions could be found previously to be treated carefully because stereo vision methods often fail in these areas.

Acknowledgment

The authors would like to thank to IF Goiano (Instituto Federal Goiano) and to CNPQ (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for financial support.

References

[1] S. Mattoccia, S. Giardino, and A. Gambini, "Accurate and efficient cost aggregation strategy for stereo correspondence based on approximated joint bilateral filtering," in *Asian Conference on Computer Vision*. Springer, 2009, pp. 371–380.

[2] R. A. Hamzah and H. Ibrahim, "Literature survey on stereo vision disparity map algorithms," *Journal of Sensors*, 2016.

[3] X. Yang, X. Chen, and J. Xi, "Block based dense stereo matching using adaptive cost aggregation and limited disparity estimation," in *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, May 2017, pp. 1–6.

[4] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, no. 1-3, pp. 7–42, 2002.

[5] N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, "Review of stereo vision algorithms: from software to hardware," *International Journal of Optomechatronics*, vol. 2, no. 4, pp. 435–462, 2008.

[6] D. Kumari and K. Kaur, "A survey on stereo matching techniques for 3d vision in image processing," *Int. J. Eng. Manuf*, vol. 4, pp. 40–49, 2016.

[7] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, Feb 2013.

[8] A. Hosni, M. Bleyer, and M. Gelautz, "Secrets of adaptive support weight techniques for local stereo matching," *Computer Vision and Image Understanding*, vol. 117, no. 6, pp. 620–632, 2013.

[9] M. Gerrits and P. Bekaert, "Local stereo matching with segmentation-based outlier rejection," in *Computer and Robot Vision, 2006. The 3rd Canadian Conference on*. IEEE, 2006, pp. 66–66.

[10] H. Hirschmuller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.

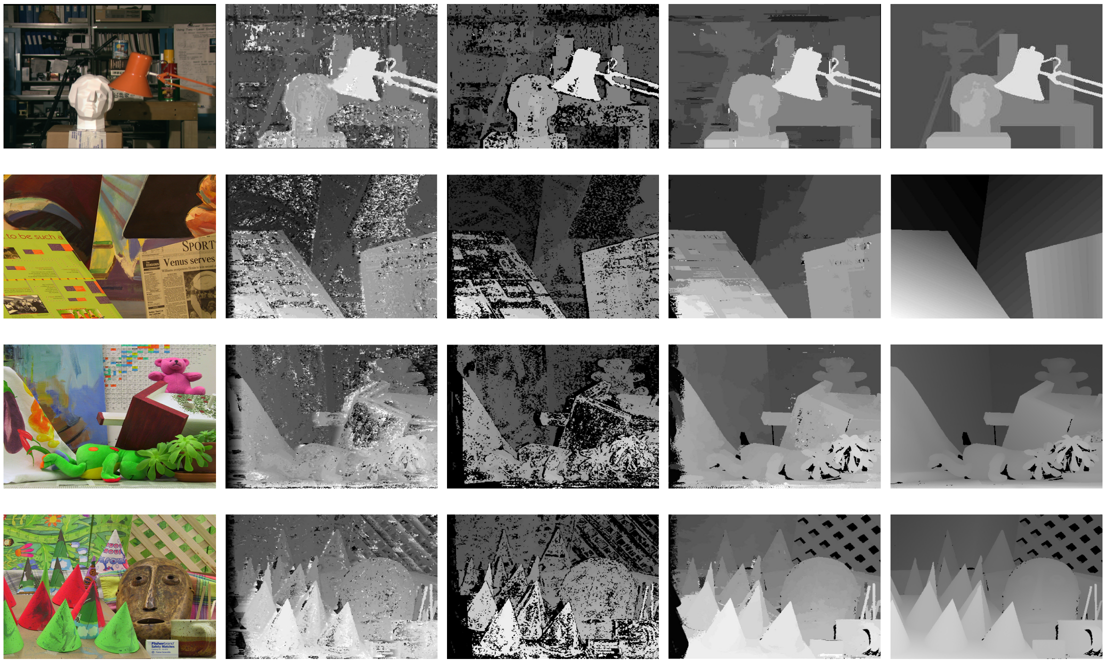
[11] H. Hirschmüller, P. R. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," *International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 229–246, 2002.

[12] Y. Wang, F. Zhong, Q. Peng, and X. Qin, "Depth map enhancement based on color and depth consistency," *The Visual Computer*, vol. 30, no. 10, pp. 1157–1168, 2014.

[13] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 30, no. 2, pp. 328–341, 2008.

TABLE I: Accuracy evaluation (in boldface, minimum bad pixel error).

	Tsukuba			Venus			Teddy			Cones		
	NOCC	ALL	DISC	NOCC	ALL	DISC	NOCC	ALL	DISC	NOCC	ALL	DISC
FW	18.98	20.81	20.05	34.78	35.86	33.81	37.32	43.79	40.75	30.01	37.95	33.68
FW + SCC	3.82	4.41	**7.44**	**10.49**	**11.12**	15.79	**13.93**	**19.96**	**19.87**	7.52	15.56	**14.15**
FW + SCCNoSpatial	**2.92**	**3.54**	9.38	10.90	11.38	**15.50**	15.17	21.26	21.72	**7.35**	**15.15**	14.24

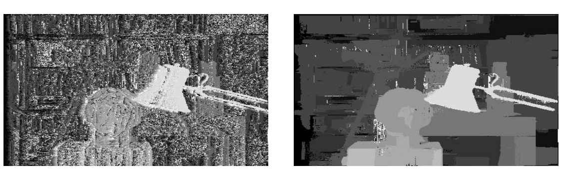


(a) (b) (c) (d) (e)

Fig. 5: Experimental results from proposed method to original images (a), tsukuba, venus, teddy and cones, respectively. Raw disparity maps in (b), unknown disparities identified after Eq. 2 in (c), SCC final results in (d) and ground truth maps in (e).

TABLE II: Pixel based technique evaluation.

	Tsukuba		
	NOCC	ALL	DISC
FW	46.47	47.63	39.63
FW + SCC	10.26	10.87	11.59
FW + SCCNoSpatial	9.08	9.60	15.35



(a) (b)

Fig. 6: Result from pixel based technique: (a) the input raw map and (b) SCC output.

[14] S. Mattoccia, "A locally global approach to stereo correspondence," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 1763–1770.

[15] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.

[16] N. Ma, Y. Men, C. Men, and X. Li, "Accurate dense stereo matching based on image segmentation using an adaptive multi-cost approach," *Symmetry*, vol. 8, no. 12, p. 159, 2016.

[17] F. Tombari, S. Mattoccia, and L. Di Stefano, "Segmentation-based adaptive support for accurate stereo correspondence," *Advances in Image and Video Technology*, pp. 427–438, 2007.

[18] K.-J. Yoon and I.-S. Kweon, "Locally adaptive support-weight approach for visual correspondence search," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, June 2005, pp. 924–931 vol. 2.

[19] D. Chen, M. Ardabilian, and L. Chen, "A fast trilateral filter-based adaptive support weight method for stereo matching," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 730–743, May 2015.

[20] G. T. Laureano and M. S. V. de Paiva, "Disparities maps generation employing multi-resolution analysis and perceptual grouping," in *2008 First Workshops on Image Processing Theory, Tools and Applications*, Nov 2008, pp. 1–6.

[21] D. Scharstein and R. Szeliski, "Middlebury stereo visions." [Online]. Available: http://vision.middlebury.edu/stereo/

[22] G. F. et al, "Integral images for block matching (demo)." [Online]. Available: http://demo.ipol.im/demo/57/