HYBRID-BASED DENSE STEREO MATCHING

T. Y. Chuang [a, *], H. W. Ting [a], J. J. Jaw [a]

[a] Dept. of Civil Engineering, National Taiwan University, Taiwan - (tzuyichuang, r01521113, jejaw)@ntu.edu.tw

Commission III, WG III/3

KEY WORDS: Edge Constraint, Shape-adaptive Cross-based Stereo Matching, Semi-Global Matching, Penalty Estimation

ABSTRACT:

Stereo matching generating accurate and dense disparity maps is an indispensable technique for 3D exploitation of imagery in the fields of Computer vision and Photogrammetry. Although numerous solutions and advances have been proposed in the literature, occlusions, disparity discontinuities, sparse texture, image distortion, and illumination changes still lead to problematic issues and await better treatment. In this paper, a hybrid-based method based on semi-global matching is presented to tackle the challenges on dense stereo matching. To ease the sensitiveness of SGM cost aggregation towards penalty parameters, a formal way to provide proper penalty estimates is proposed. To this end, the study manipulates a shape-adaptive cross-based matching with an edge constraint to generate an initial disparity map for penalty estimation. Image edges, indicating the potential locations of occlusions as well as disparity discontinuities, are approved by the edge drawing algorithm to ensure the local support regions not to cover significant disparity changes. Besides, an additional penalty parameter P_e is imposed onto the energy function of SGM cost aggregation to specifically handle edge pixels. Furthermore, the final disparities of edge pixels are found by weighting both values derived from the SGM cost aggregation and the U-SURF matching, providing more reliable estimates at disparity discontinuity areas. Evaluations on Middlebury stereo benchmarks demonstrate satisfactory performance and reveal the potency of the hybrid-based dense stereo matching method.

1. INTRODUCTION

Stereo matching is the problem of recovering corresponding points from different image views, and is one of the indispensable ingredients for 3D exploitation of imagery in the fields of Photogrammetry and Computer vision. Many different stereo vision algorithms have been proposed. An overview of state-of-the-art methods can be found in the Middlebury Stereo benchmark (Scharstein and Szeliski, 2002). Current stereo matching techniques can be categorized into pixel-wise, local and global algorithms. Pixel-wise matching methods simply take an optimal disparity for each image component. (Birchfield and Tomasi, 1999). Local algorithms aggregate the support from the neighboring pixels in a constrained region to determine the disparity without considering neighboring connections. These methods may suffer from a lack of smoothness and fail when support regions of pixels contain repetitive patterns, disparity discontinuities, and occlusions. Studies, such as adaptive windows (Kanade and Okutomi, 1994; Scharstein and Szeliski, 2002), multiple windows (Hirschmuller et al., 2002), and image segmentation (Gerrits and Bekaert, 2006), tended to adapt the shape and size of the support region or adjust the support weights of the region pixels near depth discontinuities to improve disparity estimates. The adaptive-window method finds an optimal window based on the local variation of intensity and disparity, while the multiple-window method calculates the correlation with nine pre-defined windows and selects the disparity with the smallest matching cost. These methods, nevertheless, are not suitable for arbitrarily shaped depth discontinuities due to utilizing a rectangular window. Besides, studies, such as Xu et al. (2002) who resolved adaptive support weights by radial computations; Yoon and Kweon (2006) who assigned a support weight to the pixel in a support region based on color similarity and geometric proximity; Zhang et al. (2009) who used color similarity and connectivity constraints to construct an upright cross local support skeleton for each anchor pixel, and dynamically built a shape-adaptive full support region, can also be found in the literature. Typically, local methods require less computation, and consequently less accuracy can be attained, as compared to employing global methods. Also, deciding an appropriate support region of each pixel remains an inevitable challenge for local stereo matching algorithms. On the other hand, global algorithms seek optimal pixel matches by minimizing a global energy function which consists of a data term and a smooth term to penalize inconsistent solutions and to enforce the piecewise smoothing assumption, respectively. The data term represents the coherence of a region often measured from image color, luminance, and texture. The smoothness term assigns a large penalty to those neighboring pixels conveying different disparity values, then the similar pixels can be merged. Various optimization techniques, such as graph cuts (Boykov et al., 2001; Wang et. al, 2013), belief propagation (BP) (Sun et al., 2003; Felzenszwalb and Huttenlocher, 2006; Klaus et al., 2006) and dynamic programming (DP) (Ohta and Kanade, 1985; Gong and Yang, 2003; Torr and Criminisi, 2004), are often used to determine the local minimum of the energy function. Compared to local methods, global methods render better quality of the estimated depth, but involve high computational complexity. Thus, it makes them inapplicable for near or real-time applications. In addition, Hirschmuller (2008) proposed a popular semi-global matching (SGM) to achieve high precision depth estimation, which joints the advantages of both global and local algorithms and is capable of real-time demands. SGM approximates a global optimization by combining several local optimization steps. The energy function of SGM comprises a data term, a smoothness term for slight changes in disparity, and a larger smoothness term for depth discontinuities with significant disparity changes. Klette et al. (2011) gave tests on different stereo algorithms, including variants of DP, BP, and SGM, among which they suggested that SGM can potentially deal with

* Corresponding author

scenes of high-depth complexities and is the most promising. Numerous variants of SGM, such as PlaneFitSGM (Humenberger et al., 2010), wSGM (Spangenberg et al., 2013), SGMDDW (Michael et al., 2013), rSGM (Spangenberg et al., 2014), iSGM (Hermann and Klette, 2013), eSGM (Hirschmuller, 2012), and real-time implementation (Banz et al., 2010), have been proposed with respect to advancing the accuracy and computational complexity. Banz et al. (2012) evaluated four different penalty functions under two matching cost calculation manners, namely the census and rank transformation (Zabih and Woodfill, 1994), and concluded that the linear and inversely proportional penalty functions significantly outperform the constant and variance-based ones. Besides, for all penalty functions, using the census transform instead of the rank transform exhibits better disparity maps with less edge blurring because the census transform retains spatial information. Furthermore, image features are frequently used to assist in stereo matching as well (Sadeghi et al., 2008). Yang and Wang (2015) purported to improve an adaptive support weight approach by incorporating Canny edges. Poddar et al., (2015) employed image segmentation and feature point segment matching techniques for dense disparity estimation. Xiao et al. (2013) employed ground control points into the energy function of SGM (GCP-SGM) as soft constraints for aerial image applications. Zhu et al., (2011) expended the stereo matching technique to aerial and satellite image sequences. Konrad and Lan (2000) combined the block-based disparity estimation with feature point matching.

In this paper, a hybrid-based approach is proposed to perform disparity estimation from rectified/epipolar stereo pairs. As suggested in Tian et al. (2013), matching cost is determined through the census transformation (CT) with the absolute differences (AD). A gradual strategy, involving edge constraints and penalty estimation, is presented to alleviate the sensitiveness towards the penalty parameters in cost aggregation step, and thus more accurate disparity estimation can be achieved. Image edges, which typically indicate depth discontinuities, are detected by the edge drawing (ED) algorithm (Akinlar and Topal, 2011) and treated as the edge constraint. Moreover, to shrink the potential disparity errors and to enhance the quality of depth estimation at image edges (object boundaries), this study manipulates the U-SURF descriptor (Bay et al., 2008) to simultaneously recover the edge disparities, and weights both values derived from the stereo and U-SURF matching to find the final disparities at edge pixels. Notably, in this paper no post-interpolation process is applied so as to properly evaluate the estimated results.

The rest of the paper is organized as follows: Section 2 briefly reviews the relevant techniques and introduces the methodology to address this work. Afterwards, experimental outcomes are exhibited in Section 3. Finally, conclusions are drawn in Section 4.

2. HYBRID-BASED STEREO MATCHING

The study conceptually integrates algorithms of shape-adaptive cross-based local stereo matching and semi-global matching involved with a pixel-wise edge constraint to improve the quality of disparity estimation. The proposed method starts by categorizing image content into edge and interior pixels. For disparity initialization, the edge pixels are introduced into a U-SURF matching process while the interior pixels are submitted to a gradual stereo matching process. In addition, a reasonable penalty estimation method as well as an image edge constraint are imposed on the stereo matching procedure to enhance the matching performance, and the edge disparity results derived from the two matching process are weighted for better estimation in depth discontinuity areas. Figure 1 shows the block diagram of the proposed working scheme.

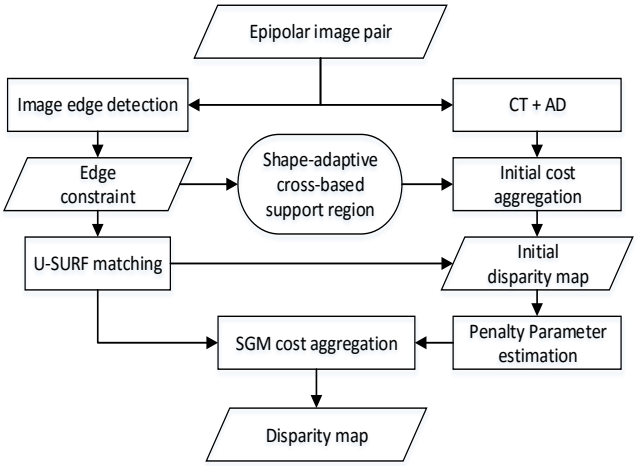


Figure 1. The block diagram of the proposed method.

2.1 U-SURF matching

Edge pixels in each image are approved by ED which takes no parameter tuning and is capable of achieving real-time processing. The study constructs U-SURF descriptors for each edge pixel and performs 1-D matching along its epipolar line within the disparity search range as illustrated in Figure 2. The U-SURF descriptor provides a unique and robust description for the intensity distribution of surrounding pixels. A square region is extracted centered on the edge pixel and split up into smaller 4x4 square sub-regions from which the Haar responses weighted with a Gaussian are extracted. The best candidate match is found through the best-bin-first search (Beis and Lowe, 1997) identifying the nearest neighbor which is defined as the one with the minimum Euclidean distance from the descriptor vector. The probability of a correct match is determined by considering the ratio of distance from the closest neighbor to the distance of the second closest. Subsequently, a quasi-RANSAC approach (Stamatopoulos et al., 2012) for outlier removal is applied. Finally, the preliminary disparities of edge pixels are resolved.

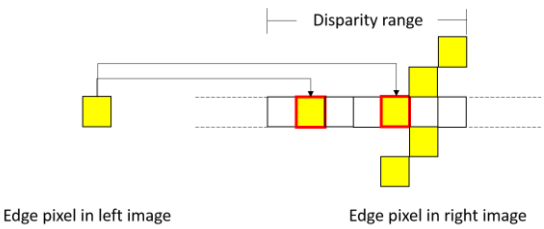


Figure 2. Illustration of the edge pixel matching.

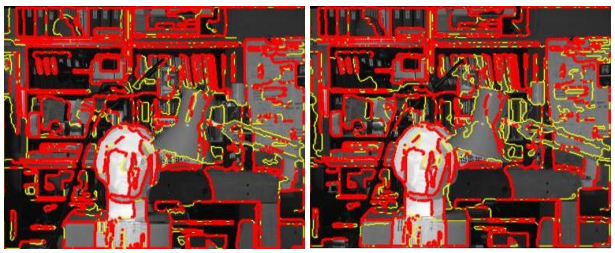
Notably, in order to prevent outliers, strict thresholds are applied to the matching ratio test and the quasi-RANSAC process. These thresholds can be further fine-tuned considering the percentage of total matched edge pixels:

$$P_{edge\ pixel} = N_{matched}/N_{detected} \quad (1)$$

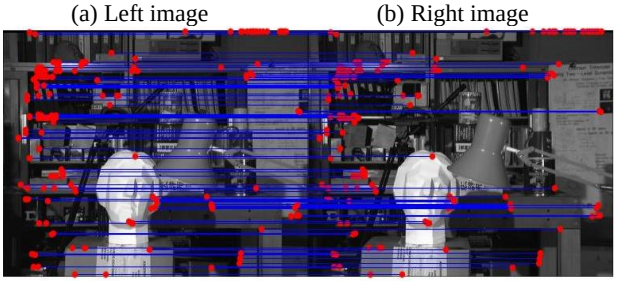
where $P_{edge\ pixel}$ = percentage of matched edge pixels
$N_{matched}$ = number of matched edge pixels
$N_{detected}$ = number of detected edge pixels

In this paper, we determine the threshold for ratio test as 0.6 and one-pixel allowance for outlier removal, an average matching

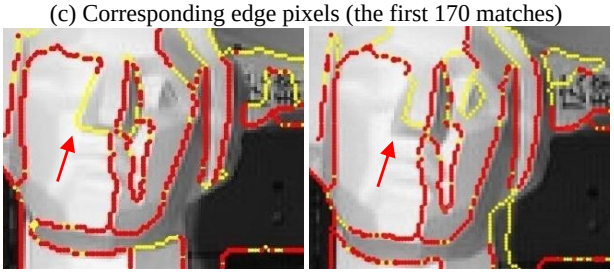
rate of 64.3% can be obtained, as portrayed in Figures 3(a) and 3(b). Figure 3(c) demonstrates the matching results of the first 170 corresponding edge pixels along their epipolar lines. On the other hand, missed matched edge pixels largely result from the occlusions and the dissimilar edge detection results of each image, as pointed out in Figures 3(d) and 3(e). The yellow dots indicate the detected edges and the red dots present the matched edge pixels in Figure 3.



(a) Left image (b) Right image



(c) Corresponding edge pixels (the first 170 matches)



(d) Enlarged look of left image (e) Enlarged look of right image

Figure 3. Demonstration of the edge pixel matching.

The disparities of those unmatched edge pixels are then explored with interior pixels as follows.

2.2 Gradual stereo matching

2.2.1 Matching cost: For accurate stereo matching, it is important to decide an appropriate metric function of matching cost for each pixel. In light of the evaluations in Hirschmuller and Scharstein (2009), the census transformation (Zabih and Woodfill 1994) proved to have the most balanced performance and better adaptability to intensity distortion caused by illuminating variation. The matching cost is decided based on the relationship of regions between two pixels instead of relying directly on the intensity values. Nevertheless, the census transformation is inadequate to handle image areas with little texture as well as repeated patterns and has a lower tolerance against image noise. Tian et al. (2013) suggested the absolute difference function with better matching effects in sparse textured areas and thus can be used to compensate for the deficiency of the census transformation. Therefore, the combination of these two matching energy functions is preferred in this paper. The integration of CT and AD matching energy functions can be expressed as (Tian et al., 2013):

$$C(p,d) = \rho(D_{census}(p,d), \lambda_{census}) + \rho(D_{AD}(p,d), \lambda_{AD}) \quad (2)$$

where d = the presupposed disparity value
D_{census} = the matching cost of CT
D_{AD} = the matching cost of AD

And the normalized function $\rho(D,\lambda)$ is formulized as:

$$C_{census}(I(p)) = \otimes_{p' \in N(p)} \xi(p,q), \xi(p,q) \begin{cases} 1 & if\ I(p) < I(q) \\ 0 & otherwise \end{cases}$$
$$D_{census}(p,d) = D_H(C_{census}(I_L(x,y)), C_{census}(I_R(x-d,y)))$$
$$D_{AD}(p,d) = \min\{|I_L(x,y) - I_R(x-d,y)|, \tau\}$$
$$\rho(D,\lambda) = 1 - \exp(-D/\lambda) \quad (3)$$

The $I_L(x,y)$ and $I_R(x,y)$ indicate gray values of corresponding pixels in the left and right images, respectively; $I(p)$ is the luminance value of pixel of p ; $N(p)$ is the p -centered transforming window; q indicates neighboring pixels. λ_{census}, λ_{AD}, and τ are prior parameters. Normalized function avoids the matching cost overly leaning to one certain type. λ controls the weights of the two matching cost, which assists in combining the advantages of CT and AD for a better adaptability to intensity distortion and noise. Details of Equations 2 and 3 can be referred to Tian et al. (2013).

2.2.2 Initial cost aggregation: SGM cost aggregation aims to minimize several 1-D energy functions for a global 2-D energy minimization problem. Typically, a 1-D energy function $E(D)$ comprising a data term and a smooth term can be formulated as (Hirschmuller and Scharstein, 2009):

$$E(D) = \sum_p (c(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \quad (4)$$

The data term calculates the sum of a pixel-wise matching cost $c(p, D_p)$ for all pixels p at their disparities D_p; the smooth term involves two penalty parameters (P_1 , P_2). P_1 penalizes neighboring pixels N_p of p if their disparity difference is equal to 1; P_2 is imposed on disparity changes larger than 1. That is, P_1 permits an adaptation to slanted or curved surfaces while P_2 preserves discontinuities as they mostly coincide with intensity variations. Yet, the value of P_2 should be larger than P_1 . Basically, the quality of the disparity map significantly relies on the two empirically determined parameters for standard SGM (Hermann et al., 2009). If the value of P_2 is too small to smoothen a disparity map, most noise would remain. Conversely, the disparity map would be over smoothened with an extremely large P_2. Hirschmuller (2008) first introduced an adaptive P_2 function to penalize abrupt disparity changes according to the image content. Banz et al. (2012) evaluated three further penalty functions and concluded that the negatively and inversely proportional to the absolute luminous intensity gradient of the currently processed pixels along the path significantly outperform the variance-based approach. However, the parameters involved in each function still need to be manually tuned with a carefully determined (try and error) step in a way they must be big enough to ensure sufficiently different configuration and small enough not to miss local minima at the same time.

In current work, a proper penalty estimation and the edge constraint are introduced into the cost aggregation step to facilitate the disparity computation. The aggregation process comprising initial disparity generation, penalty estimation, and disparity determination is given as follows.

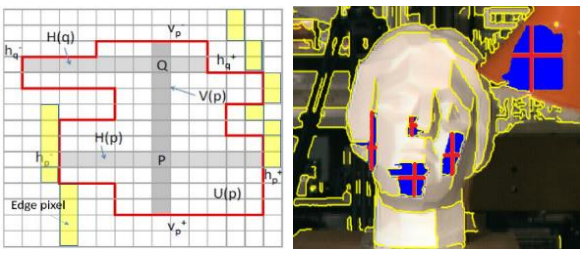
2.2.3 Initial disparity generation: The study leverages shape-adaptive cross-based matching approach (Zhang et al., 2009) with the edge constraint to generate an initial disparity map. The edge constraint enforces the adaptive support regions not to cross significant disparity changes, which is appropriate for pixels near arbitrarily shaped depth discontinuities. The adaptive cross-based algorithm is performed on the interior pixels to search their support region in four directions (upper, lower, left and right). The region stretches from the central pixel until two consecutive pixels are not consistent in color, making the maximum arm length, or touching the edge pixels. Edge pixels are excluded from this measure. The arm length r^* is defined as (Zhang et al., 2009):

$$r^* = max_{r\in[1,L]}(r \prod_{i\in[1,r]} \delta(p, p_i)) \quad (5)$$

where $p_i = (x_p - i, y_p)$, $(p, p_i) \notin$ edge pixels; L is the preset maximum arm length. $\delta(p, q_i)$ is an indicator function evaluating the color similarity between consecutive pixels based on all color bands.

$$\delta(p, p_i) = \begin{cases} 1, max_{c\in[R,G,B]}(|(I_c(p_1) - (I_c(p_2)|) \le \tau_c \\ 0, otherwise \end{cases} \quad (6)$$

where I_c is the intensity of the color band, and τ_c controls the confidence level of color similarity. Under the constraint of edges, the support regions of each interior pixel are reconstructed through whole image. The pixels $(v_p^+, v_p^-, h_p^+, h_p^-)$ in Figure 4(a) define the horizontal segment H(p) and vertical segment V(p). The expansion of the combined local cross H(p) ∪ V(p) determines the integrated aggregation region U(p). Figure 4(b) demonstrates a result of support regions with the edge constraint, where the edges confining the support regions are colored in yellow. To enhance the consistency of disparity values within the support region, an efficient OII voting technique (Zhang et al., 2009), which decomposes the matching cost aggregation into two orthogonal 1-D integration steps, is used. The disparity value has a maximum population within the support region is deemed as the result of the interior pixel. Accordingly, the disparities of edge and interior pixels are jointed to generate the initial disparity map.



(a) (b)

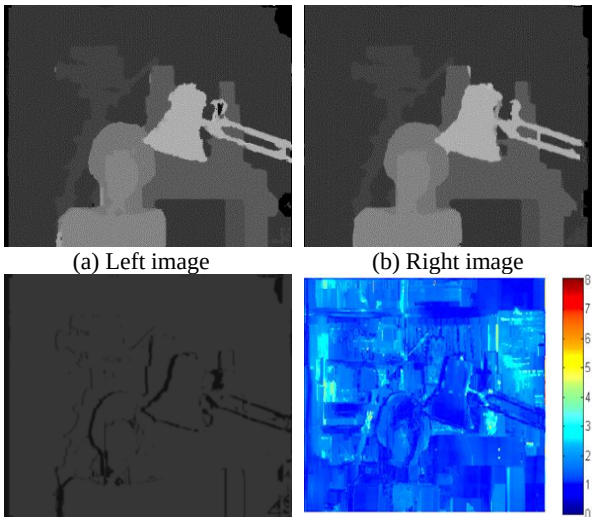
Figure 4. Demonstrations of the shape-adaptive cross-based matching cost aggregation with the edge constraint.

2.2.4 Penalty estimation: Considering the diversity of image content, it is not a trivial work to automatically arrange the penalty parameters in the cost aggregation step. This study presents a convenient way to estimate proper penalty parameters for the SGM cost aggregation. A consistency check is first done to rectify the initial disparity map, eliminating occlusions and false matches as shown in Figure 5(c). Then, a matching cost ratio (C_r) is determined for the inspected pixels.

$$Cr_i = Ms_i/Mf_i, i \in N_{ip} \quad (7)$$

where Cr_i = the ratio of matching cost
Mf_i = the smallest matching cost
Ms_i = secondary small matching cost
N_{ip} = number of inspected pixels

A larger C_r of a pixel represents a higher confidence in its disparity estimate. Thus, a map of matching cost ratio can be generated as Figure 5(d).



(a) Left image (b) Right image

(c) Errors pixels in black (d) Matching cost ratio

Figure 5. The generation of a matching cost ratio map.

Consequently, confident pixels are selected if their ratios are larger than a threshold τ_r. Therefore, P_2 is determined from the confident pixels by taking the average of the differences between the smallest matching cost and the secondary small one. P_1 is half of P_2.

$$\begin{cases} P_2 = \frac{1}{N_{cp}} \sum_{i\in N_{cp}} (Ms_i - Mf_i) \\ P_1 = P_2/2 \end{cases} \quad (8)$$

where N_{cp} = number of confident pixels

Equation 8 gives proper penalty estimates for automated parameter setting in the energy function. Still, the values can be fine-tuned if the matching result suggests this is necessary. The potency of the penalty estimation is to be demonstrated in the evaluation section.

2.2.5 Disparity determination: Standard SGM aggregates $E(D)$ along 1-D paths from eight directions toward each pixel of interest using dynamic programming. As mentioned in section 2.2.2, an over large P_2 would extremely smoothen the disparity map. Thus, a relative smaller penalty parameter should be used at image edge pixels to preserve discontinuities coincided with intensity variations. As a result, the study imposes the edge constraint and applies an additional edge penalty parameter P_e to the energy function at edge pixels as:

$$E(D) = \sum_p (c(p, D_p) + \sum_{q\in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q\in N_p} P_2 T[|D_p - D_q| > 1 \cap q \notin Eps] + \sum_{q\in N_p} P_e T[|D_p - D_q| > 1 \cap q \in Eps] \quad (9)$$

where P_e = penalty parameter for edge pixels
$P_e = (P_1 + P_2)/2$
Eps = edge pixels

The value of the edge penalty parameter P_e is defined as the average of P_1 and P_2. The disparity is then retrieved by a winner-takes-all strategy. Furthermore, we weigh the resultant disparities of edge pixels (D_{sm}) with the preliminary edge disparities (D_{um}) acquired from U-SURF matching to obtain the final disparities of edges, reading in Equation 10:

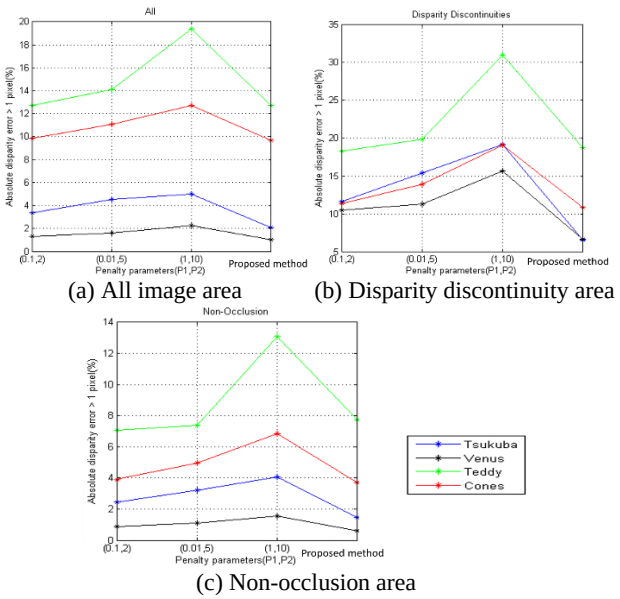
$$D_{edge} = \lambda D_{sm} + (1 - \lambda)\, D_{um} \tag{10}$$

where λ = weight value, $\lambda \in [0,1]$
D_{edge} = weighted edge disparities
D_{sm} = edge disparities derived from stereo matching
D_{um} = edge disparities derived from U-SURF matching

Finally, the disparity map of the whole image is determined.

3. EVALUATION

The experiments aim to evaluate the effectiveness of the proposed hybrid-based matching method in two phases. Test images (Tsukuba, Venus, Teddy, and Cones) of the Middlebury stereo benchmark (Scharstein and Szeliski, 2008) were utilized. Firstly, we compared the effectiveness of the penalty estimation with manual configurations of $(P_1, P_2) \in \{0.1{,}2;\ 0.01{,}5;\ 1{,}10\}$ for disparity map generation. Figure 6 shows the proportion of the disparity error of the four test image. This reveals that our method obtains the best accuracy in both disparity discontinuity and non-occlusion areas. Manual setting of P_2 performs well if carefully adjusted to the image, but quality degrades rapidly as these values are changed. In light of the resultant disparity maps (Tsukuba image, for instance) shown in Table 1, the proposed penalty estimation outperforms all the results inferred from the manual parameter setting, particularly in the region pointed out by the red arrow.



(a) All image area (b) Disparity discontinuity area

(c) Non-occlusion area

Figure 6. The proportion of the disparity error (> 1pixel).

Tsukuba	
$(P_1, P_2) = (0.01{,}5)$	$(P_1, P_2) = (0.1{,}2)$
$(P_1, P_2) = (1{,}10)$	Proposed method

Table 1. Results of the disparity map.

The study imposes the constrained shape-adaptive cross-based local method on the SGM to alleviate the sensitiveness towards the penalty parameters. Thus, we respectively jointed three additional local methods, namely an adaptive support-weight approach (ASW) (Yoon and Kweon, 2006), a fixed window-based approach (FW), and a cross-based method (CB) (Zhang et al., 2009), with the SGM cost aggregation to evaluate the virtues of the proposed method. Table 2 shows the proportion of the disparity error of each set of the stereo matching methods. Table 3 demonstrates the resultant disparity maps of the Tsukuba image, and the edge disparities derived from the U-SURF matching.

Test image	Tsukuba	Venus	Teddy	Cones
	Total (%)			
SGM	6.31	11.8	17.7	12.5
FW+SGM	3.67	3.09	16.5	11.1
ASW+SGM	3.26	2.82	15.9	10.8
CB+SGM	2.05	1.04	12.7	9.68
Proposed method	1.94	1.01	12.8	9.24
	Non-occlusion (%)			
SGM	4.80	10.7	10.9	5.12
FW+SGM	2.49	2.09	10.2	5.52
ASW+SGM	2.27	1.93	9.42	4.14
CB+SGM	1.48	0.62	7.74	3.17
Proposed method	1.44	0.59	7.71	3.13
	Disparity discontinuity (%)			
SGM	15.7	31.4	23.9	13.9
FW+SGM	12.6	21.8	24.3	13.3
ASW+SGM	11.1	20.6	22.7	11.9
CB+SGM	6.52	6.64	18.7	10.8
Proposed method	4.63	5.83	15.6	8.9

Table 2. The proportion of the disparity error (> 1 pixel).

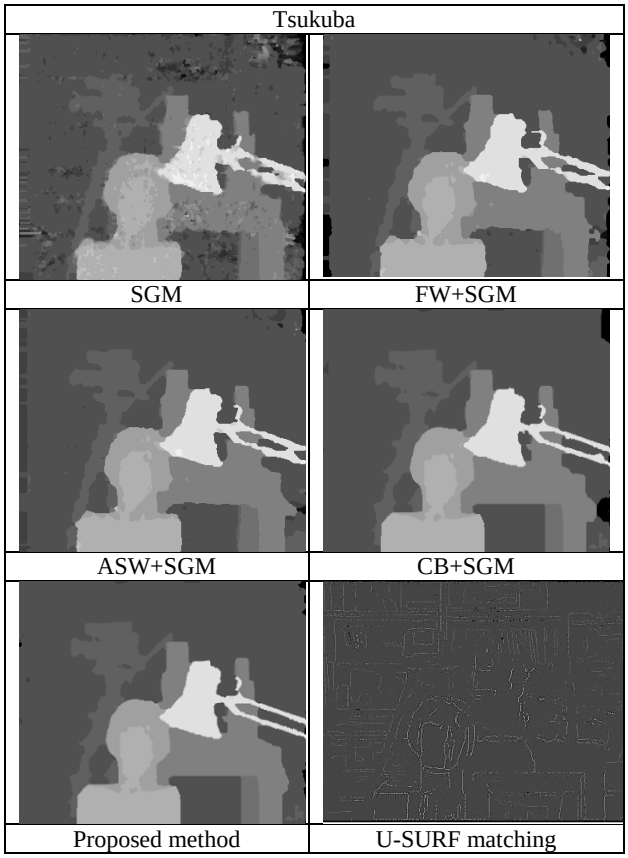


Table 3. The resultant disparity maps.

In light of the Table 2 and Table 3, SGM results in a less smooth disparity map because its cost aggregation step is susceptible to the values of the penalty parameters. Noise is effectively reduced while the local methods are incorporated into the SGM. This proves the strategy that brings forth an initial disparity map to facilitate the SGM cost aggregation is practically feasible. Moreover, among these approaches, the proposed method achieves the best accuracy for disparity estimation. Because of the impression of the edge constraint, the disparity errors of our method are relatively small in particular close to the discontinuity area. These results also suggest that the more accurate initial disparity is provided, the higher quality of the disparity estimation can be achieved.

4. CONCLUSIONS AND FUTURE WORK

In this paper, the effectiveness of the proposed hybrid-based stereo matching approach has been verified with the Middlebury stereo benchmark. The study presents the shape-adaptive cross-based matching approach with the edge constraint to generate an initial disparity map for the penalty estimation. As a consequence, the sensibility of the SGM cost aggregation towards the penalty parameters can be alleviated. In summary, the study imposes the edge constraint onto the energy function of SGM and integrates the disparities of edge pixels, derived from U-SURF matching to improve the accuracy in disparity discontinuity regions. Moreover, the optimization in reducing computational complexity and further assessment will be studied in future work.

ACKNOWLEDGEMENTS

The research was supported by the Ministry of Science and Technology, Taiwan (104-2119-M-002-025). The authors acknowledge the anonymous reviewers for their valued comments and suggestions.

REFERENCES

Akinlar, C., and Topal, C., 2011. EDLines: A real-time line segment detector with a false detection control, *Pattern Recognition Letters*, 32(13), pp. 1633-1642.

Banz, C., Hesselbarth, S., Flatt, H., Blume, H., and Pirsch, P., 2010. Real-time stereo vision system using semi-global matching disparity estimation. In: *International Conference on Embedded Computer Systems*, pp. 93-101.

Banz, C., Pirsch, P., and Blume, H., 2012. Evaluation of penalty functions for semi-global matching cost aggregation. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIX, part B3, pp. 1-6.

Bay, H., Ess, A., Tuytelaars, T., and Gool, L.V., 2008. SURF: Speeded Up Robust Features. *Computer Vision and Image Understanding*, 110(3), pp. 346-359.

Beis, J., and Lowe, D.G., 1997. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. In: *Conference on Computer Vision and Pattern Recognition*, pp. 1000-1006.

Birchfield, S., and C. Tomasi, 1999. Depth discontinuities by pixel-to-pixel stereo. *International Journal of Computer Vision*, 35(3), pp. 269-293.

Boykov, Y., Veksler, O., and Zabih, R., 2001. Fast approximate energy minimization via graph cuts, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11), pp. 1222-1239.

Felzenszwalb, P.F., and Huttenlocher, D.P., 2006. Efficient belief propagation for early vision. *International Journal of Computer Vision*, 70(1), pp 41-54.

Gerrits, M., and Bekaert, P, 2006. Local stereo matching with segmentation-based outlier rejection, In: *Proceedings of the 3rd Canadian Conference on Computer and Robot Vision*, pp. 66.

Gong, M., and Yang, Y., 2003. Fast stereo matching using reliability-based dynamic programming and consistency constrains, In: *Proceedings of IEEE Conference Computer Vision*, pp. 610-617.

Hermann, S., Klette, R., and Destefanis, E., 2009. Inclusion of a second-order prior into semi-global matching. In: *Proceedings of Pacific Rim Symposium Image Video Technology*, Vol. 5414, pp. 633-644.

Hermann, S., and Klette, R., 2013. Iterative semi-global matching for robust driver assistance systems. In: *Proceedings of Asian Conference on Computer Vision*, pp. 465-478.

Hirschmuller, H., Innocent, P., and Garibaldi, J., 2002. Real-time correlation-based stereo vision with reduced border errors. *International Journal of Computer Vision*, 47(1), pp. 1-3.

Hirschmuller, H., 2008. Stereo processing by semi global matching and mutual information. *Pattern Analysis and Machine Learning*, 30(2), pp. 328-341.

Hirschmuller, H., and Scharstein, D., 2009. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9), pp. 1582-1599.

Hirschmuller, H., Buder, M., and Ernst, I., 2012. Memory efficient semi-global matching. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. I-3, pp. 371-376.

Humenberger, M., Engelke, T., and Kubinger, W., 2010. A census-based stereo vision algorithm using modified semi-global matching and plane fitting to improve matching quality. In: *Proceedings of Computer Vision and Pattern Recognition - Workshops*, pp. 77-84.

Kanade, T., and Okutomi, M., 1994. A stereo matching algorithm with an adaptive window: theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9), pp. 920-932.

Klaus, A., Sormann, M., and Karner, K., 2006. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In: *International Conference on Pattern Recognition*, Vol. 3, pp.15-18.

Klette, R., Kruger, N., Vaudrey, T., Pauwels, K., van Hulle, M., Morales, S., Kandil, F., Haeusler, R., Pugeault, N., Rabe, C., and Lappe, M., 2011. Performance of correspondence algorithms in vision-based driver assistance using an online image sequence database. *IEEE Transactions on Vehicular Technology*, 60(5), pp. 2012-2026.

Konrad, J., and Lan, Z.D., 2000. Dense disparity estimation from feature correspondences. In: *Proceedings of the Society of Photographic Instrumentation Engineers*, Vol. 3957, pp. 90-101.

Michael, M., Salmen, J., Stallkamp, J., and Schlipsing, M., 2013. Real-time stereo vision: optimizing semi-global matching. In: *Intelligent Vehicles Symposium*, pp. 1197-1202.

Ohta, Y., and Kanade, T., 1985. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(2), pp. 139-154.

Poddar, S., Sahu, H., Bangale, M.R., Kumar, V., and kumar, A., 2015. Feature based dense disparity estimation. In: *International Conference on Industrial Instrumentation and Control*, pp. 950 - 955.

Sadeghi, H., Moallem, P., and Monadjemi, S.A., 2008. Feature based dense stereo matching using dynamic programming and color. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 2(6), pp. 2128-2135.

Scharstein, D., and Szeliski, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, *International Journal of Computer Vision*, 47(1), pp. 7-42.

Scharstein, D., and Szeliski, R., 2008. Middlebury Stereo Vision Page. http://vision.middlebury.edu/stereo/ (03 Mar. 2016)

Spangenberg, R., Langner, T., and Rojas, R., 2013. Weighted semi-global matching and center-symmetric census transform for robust driver assistance. Lecture Notes in Computer Science, Vol. 8048, PART 2, pp. 34-41.

Spangenberg, R., Langner, T., Adfeldt, S., and Rojas, R., 2014. Large scale semi-global matching on the CPU. In *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 195-201.

Stamatopoulos, C., Chuang, T.Y., Fraser, C.S., and Lu, Y.Y., 2012. Fully automated image orientation in the absence of targets. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIX, B5, pp. 303-308.

Sun, J., Zheng, N.N., and Shum, H.Y., 2003. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7), pp. 787-800.

Tian, Z., Xu, C., and Huang, X., 2013. Real-time stereo matching using adaptive window based disparity refinement. In: *Proceedings of the International Conference on Multimedia and Human Computer Interaction*.

Torr, P.H.S, and Criminisib, A., 2004. Dense stereo using pivoted dynamic programming. *Image and Vision Computing*, 22(10), pp. 795-806.

Wang, Y.C., Tung, C.P., and Chung, P.C., 2013. Efficient disparity estimation using hierarchical bilateral disparity structure based graph cut algorithm with a foreground boundary refinement mechanism. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(5), pp. 784-801.

Xiao, X., Guo, B., Pan, F., and Shi, Y., 2013. Stereo matching with weighted feature constraints for aerial images. In: *International Conference on Image and Graphics*, pp. 562-567.

Xu, Y., Wang, D., Feng, T., and Shum, Y.H., 2002. Stereo computation using radial adaptive windows. In: *Proceedings of Pattern Recognition*, Vol. 3, pp. 595-598.

Yang, S.H., and Wang, B., 2015. Improved adaptive-support stereo matching using local features. *Journal of Information Science and Engineering*, 31(5), pp. 1557-1575.

Yoon, K., and Kweon, S., 2006. Adaptive support-weight approach for correspondence search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4), pp. 650-656.

Zabih, R., and Woodfill, J., 1994. Non-parametric local transforms for computing visual correspondence. In: *Proceedings of European Conference on Computer Vision*, pp. 151-158.

Zhang, K., Lu, J., and Lafruit, G., 2009. Cross-based local stereo matching using orthogonal integral images. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(7), pp. 1073-1079.

Zhu, K., D'Angelo, P., and Butenuth, M., 2011. A performance study on different stereo matching costs using airborne image sequences and satellite images. In: *Proceedings of the ISPRS conference on Photogrammetric image analysis*, pp. 159-170.