https://doi.org/10.2352/ISSN.2470-1173.2021.2.SDA-059

A New Hybrid Stereo Disparity Estimation Algorithm with Guided Image Filtering-Based Cost Aggregation

Hanieh Shabanian, Madhusudhanan Balasubramanian; Electrical and Computer Engineering Department, The University of Memphis; Memphis, Tennessee, USA

Abstract

Stereo matching algorithms are useful for estimating a dense depth characteristic of a scene by finding corresponding points from stereo images of the scene. Several factors such as occlusion, noise, and illumination inconsistencies in the scene affect the disparity estimates and make this process challenging. Algorithms developed to overcome these challenges can be broadly categorized as learning-based and non-learning based disparity estimation algorithms. The learning-based approaches are more accurate but computationally expensive. In contrary, non-learning based algorithms are widely used and are computationally efficient algorithms. In this paper, we propose a new stereo matching algorithm using guided image filtering (GIF)-based cost aggregation. The main contribution of our approach is a cost calculation framework which is a hybrid of cross-correlation between stereo-image pairs and scene segmentation (HCS). The performance of our HCS technique was compared with state-of-the-art techniques using version 3 of the benchmark Middlebury dataset. Our results confirm the effective performance of the HCS technique.

1. Introductions and Related Works

Stereo matching techniques estimate a disparity map comprised of dense pixelwise correspondences between rectified stereo-image pairs of a scene with only horizontal disparity present. The horizontal disparity in stereo vision refers to an apparent shift between the left- and right-channel images in a stereo pair due to the underlying geometry of the stereo vision setup. The generated disparity map and perceived depth of the images are inversely proportional. Hence, for stereo cameras with parallel optical axes, focal length f, baseline B, and disparity d, the perceived depth is calculated by triangulation as $(B * f) \,/\, d$. Stereo vision is useful in many fields such as for robotic navigation [1], 3D surface reconstruction [2], face recognition [3], and autonomous driving [4]. Stereo matching is a challenging problem because of the presence of occlusion, distortion, noise, and ambiguous nature of the imaging environment such as presence of large homogeneous regions in the scene.

Stereo matching algorithms can be broadly classified as global and local algorithms [5]. In the global stereo matching algorithms, the disparity map is obtained by optimizing an energy or cost as a function of data and spatial smoothness of the solution (e.g. graph cut [6], belief propagation [7]). Though this category of the algorithms provides an accurate disparity estimate, they are computationally expensive. In contrast to the energy-based approaches, the local methods utilize a cost function defined based on corresponding spatial locations with a support window in the stereo-image pairs (e.g. correlation, sum of absolute differences). Hence, local methods have low complexity and are also suitable for real-world applications.

Most of the local stereo matching algorithms follow the same framework comprised of an initial matching cost calculation, cost aggregation, optimal disparity selection, and final disparity refinement steps. Among these steps, the initial matching cost computation step is important for obtaining an accurate stereo disparity map and the cost aggregation step is essential for obtaining more robust results in the presence of noise, scene occlusion, and pixel homogeneity. Commonly used data cost are sum of absolute difference (SAD) [8], gradient similarity [9], sum of squared difference (SSD) [10], non-parametric transforms [11], and normalized cross-correlation [12].

Recently, several cost aggregation strategies have been introduced to overcome the above-mentioned challenges in stereo disparity estimation. He et al. [13] developed a superior edge-preserving filter called guided image filtering (GIF). Hosni et al. [14] extended GIF for faster computation and utilized the GIF for cost aggregation by fast cost-volume filtering (FCVF). In [15] Hong et al., GIF computational time was further improved by computing linear coefficients using sub-sampled slices of the cost volume. Li et al. [16] took advantage of global and local smoothing filters and proposed weighted guided image filtering (WGIF) to achieve better visual quality. Using this modified regularized filter, Hone et al. [17] achieved a limited improvement in performance by applying WGIF to their pixelwise matching cost function. Based on WGIF, Zhu et al. [18] proposed a stereo matching algorithm with a modified census transform and constructed an adaptive rectangular support window to perform the cost aggregation. Yoon et al. [19] introduced the use of adaptive support weight (ASW) in local methods which was utilized and improved in various frameworks such as from Yang et al. [20], Zhu et al. [18]. Recently, using GIF or its extension, Hong and Kim [21], Zhu and Chang [22], and Han et al. [23] reported significant improvements in the disparity estimates.

In this paper, we propose a new stereo matching algorithm using GIF based cost aggregation. The main contribution of our method lies in the cost calculation framework using a hybrid of a priori scene segmentation and cross-correlation between stereo images (HCS) to generate an initial disparity map. We used the Middlebury dataset to test our proposed approach and extensively compared it with recent state-of-the-art approaches based on the Middlebury evaluation stereo images version 3.0.

The rest of the paper is organized as follows. Section 2 provides a detailed description of the proposed stereo matching algorithm. Experimental results and discussion are presented in Section 3 and our conclusion is presented in Section 4.

2. Proposed Algorithm

The proposed stereo-matching algorithm consists of five main steps namely, preprocessing, initial cost computation, cost aggregation, disparity computation, and post-processing.

2.A. Preprocessing

We used homomorphic filtering to correct for any illumination differences in the rectified image pair before disparity estimation [24].

In homomorphic filtering, an image I is first decomposed into an illumination component l and a reflectance component r as

$$I(x,y) = l(x,y)\ r(x,y) \quad (1)$$

The reflectance component is related to the inherent optical properties of objects that comprise the scene while the illumination component arises from the scene lighting condition and camera orientation. Therefore, to correct for illumination differences, the illumination component is minimized, and the reflectance component of the image is retained by the homomorphic filter. In general, because illumination variation across a scene is gradual, a low-pass filter is used to estimate the illumination component in the image. In log-domain,

$$\log(I(x,y)) = \log(l(x,y)) + \log(r(x,y)) \quad (2)$$

Therefore, we estimated the illumination-corrected images by subtracting the illumination component of an image from each of the respective images in the log-domain. Let, $\hat{r}_l$ and $\hat{r}_r$ are the illumination corrected (reflectance dominant) stereo image pair. In our proposed hybrid technique, these illumination-corrected stereo image pairs were used for disparity estimation.

2.B. Initial Cost Computation using Our Hybrid Technique

Based on the *epipolar constraint* [25], location corresponding to any point in one of the images in the stereo pair is bound to lie along an *epipolar line* in the other image in the stereo pair. For example, any point in the left channel image $\hat{r}_l$ lies along a corresponding epipolar line in the right channel image $\hat{r}_r$. After stereo rectification, these epipolar lines in each of the images in the stereo pair become horizontal and are aligned with the scanlines containing the corresponding points in the other rectified image in the stereo pair [26]. With one of the images in the stereo pair as reference (e.g. left channel $\hat{r}_l$), the difference in the coordinates of the corresponding points in the stereo pair is an estimate of stereo disparity. Only horizontal disparity exists between rectified stereo images.

To minimize the computational complexity of estimating dense disparity at each of the pixel locations, we first segmented the reference image based on its regional texture characteristics using an unsupervised texture segmentation algorithm [27]. Next, we estimated a distribution of disparity estimates for each of the segmented zones in the reference image. The zonal disparity distributions were used as *a priori* disparity estimates for the respective image regions or segments to estimate dense disparities.

The unsupervised texture segmentation algorithm uses a Gabor filterbank to derive multi-resolution texture features capable of defining sharp segment boundaries. Using the *k*-means clustering algorithm, the Gabor texture features are integrated to generate an image segmentation map [27]. For generating *a priori* distribution of zonal disparities, we evaluated various feature detection techniques namely the minimum eigenvalue algorithm [28], speeded up robust features algorithm (SURF) [29], Harris corner and edge detector algorithm [30] and accelerated segment test algorithm (FAST) [31]. At sparse image locations with robust features, disparity was estimated by identifying spatial correspondences between the stereo images by feature matching. The sparse disparity estimates within each segment in the reference image were pooled to generate a zone-specific distribution of stereo disparities. Thus, all locations within a zone or segment in the reference image were assigned *a priori* mean disparity μ_d and offset disparity σ_d, where the offset disparity captures the maximum disparity deviation at any location within the zone with respect to the zonal mean disparity. We expect that this *a priori* characterization of disparity distribution in each zone can reduce the computational complexity in generating a dense disparity map.

Disparity at each of the locations in the reference image were estimated using a template matching procedure as follows. For each location (x,y) in the reference image, a template g_l centered at that location in the reference image and a horizontal image strip h_r centered at coordinate $(x+\mu_d, y)$ with a width of $2\sigma_d + 1$ from the second image were selected. Thus, the search zone for identifying disparity at each location was limited by the *a priori* zonal disparity distribution along the respective scanline. A cost volume $C(x,y,d)$ at each of the locations (x,y) in the reference image for various possible choices of disparity d was built using a normalized cross-correlation measure. For faster computation, the correlation coefficients were estimated in the Fourier domain [32]. A dense disparity map can be obtained from cost optimization as

$$\hat{d}(x,y) = \underset{d}{\operatorname{argmax}}\ C(x,y,d) \quad (3)$$

2.C. Cost Aggregation and Disparity Estimation

Prior to cost optimization, the cost volume $C(x,y,d)$ at each disparity level was smoothed using a guided image filter as in He et al [13]. A unique spatially varying filter at each of the locations (x,y) in the reference image $\hat{r}_l$ is defined with respect to the reference image $\hat{r}_l$ based on the intensity distribution within a window ω_{xy} centered at (x,y) in the reference image. Let, $|\omega_{xy}|$ be the number of locations within the window ω_{xy}; μ_{xy} be the mean illumination corrected intensity within the window; and σ_{xy} be its standard deviation. Guided filter W with a regularization parameter ϵ for location (x,y) is defined as

$$W_{xy}(i,j) = \frac{1}{|\omega_{xy}|^2}\sum_{\forall(k,l)\in\omega_{xy}}\left(1 + \frac{(\hat{r}_l(i,j)-\mu_{xy})\ (\hat{r}_l(k,l)-\mu_{xy})}{\sigma_{xy}^2+\epsilon}\right) \quad (4)$$

To derive a spatially coherent disparity map $\hat{d}(x,y)$, an aggregated cost volume $\hat{C}$ was estimated using the guided filter W_{xy} as

$$\hat{C}(x,y,d) = \sum_{\forall(\hat{x},\hat{y})} W_{xy}\ (\hat{x},\hat{y})\ C(\hat{x},\hat{y},d) \quad (5)$$

Disparity at each location (x,y) was estimated using a winner-take-all optimization strategy as

$$d(x,y) = \underset{d}{\operatorname{argmax}}\ \hat{C}(x,y,d) \quad (6)$$

2.D. Post-Processing

To identify occluded locations, a left-right consistency check was used. Locations with differing disparity estimates with left channel image as the reference image and the right channel image as the reference image were identified as occluded locations. Disparity at the occluded locations were assumed to be the same as the disparity of the nearest non-occluded location in the same scanline. A median filter was applied to obtain a final dense disparity map.

3. Experimental Results and Discussion

The proposed algorithm was implemented in MATLAB 2018a on an Intel(R) Xeon(R), CPU E3-1271 v3, 3.6 GHZ processor. We performed extensive experiments to verify the effectiveness of the proposed method for disparity map estimation. At first, we tested the proposed framework using three different datasets [33],[34],[35] of the Middlebury benchmark stereo database [36].

Then, we compared the proposed method with related state-of-the-art approaches based on the Middlebury evaluation dataset version 3.0. Disparity error maps, peak signal-to-noise ratio (PSNR), and average absolute error (Avg. err in pixels) were calculated for each of the test data for performance evaluation.

PSNR provides a measure of similarity between the estimated disparity map $\hat{d}(x,y)$ and the ground-truth disparity map $d(x,y)$ each of size $m \times n$ pixels.

$$PSNR = 10\log_{10}\frac{255^2 \times m \times n}{\sum_{\forall(x,y)}\left(\hat{d}(x,y)-d(x,y)\right)^2} \quad (7)$$

Disparity error maps were computed as location-wise difference between the estimated disparity $\hat{d}(x,y)$ and its ground-truth $d(x,y)$ as $\hat{d}(x,y) - d(x,y)$. A thresholded average disparity error metric with a threshold of T pixels of disparity difference was defined as

$$Bad = \left(\frac{1}{mn}\sum_{\forall(x,y)}\left(\left|\hat{d}(x,y) - d(x,y)\right|\right) > T\right) \times 100 \quad (8)$$

The average disparity error was quantitatively assessed at three threshold levels namely Bad1.0 with $T = 1$ pixel, Bad2.0 with $T = 2$ pixels, and Avg. err with $T = 0$ pixels.

3.A. Evaluation of the Initial Disparity on the Middlebury Benchmark

Quantitative measures of accuracy of the initial disparity map generated by our proposed hybrid method are presented in Table 1.

Table 1: Quantitative evaluation of the initial disparity map generated by our proposed hybrid method on a subset of Middlebury stereo dataset.

Metrics	Cones	Teddy	Baby1	Cloth3
Avg. err	5.88	5.18	3.91	5.30
PSNR (dB)	25.20	26.33	29.66	25.70
Bad 1.0 (%)	27.22	27.14	25.89	21.81
Bad 2.0 (%)	25.69	24.21	23.52	20.65
Run time (s)	42.50	39.68	42.94	44.21

Figure 1 shows the initial disparity maps estimated using the proposed approach on a subset of Middlebury dataset namely, Cones, Teddy, Baby1 and Cloth3 stereo pairs along with the ground truth and the disparity error map.

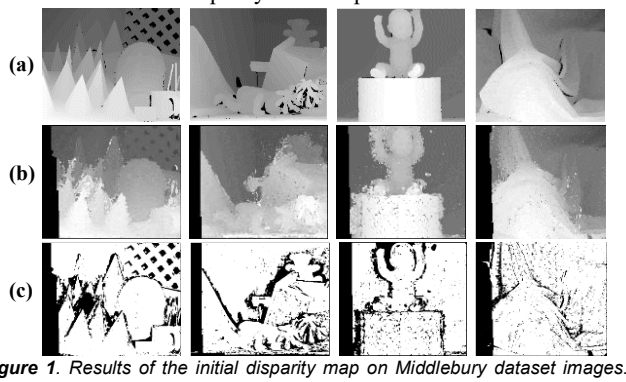


***Figure 1**. Results of the initial disparity map on Middlebury dataset images. Cones, Teddy, Baby1, and Cloth3 from left to right. (a) Ground truth, (b) the initial estimated results from the proposed approach, and (c) error map. The occluded regions are included on all the obtained results.*

3.B. Evaluation of the Cost Aggregation on the Middlebury Benchmark

In this section, we evaluate the performance of the proposed algorithm after cost aggregation using guided image filtering. The filter kernel window size was set to 7×7 pixels, and the regularization parameter ϵ initialized to 1×10^{-7}. Quantitative error metrics of PSNR, Bad1.0, Bad2.0 and Avg. err for the cost-aggregated disparity estimates are presented in Table 2.

Table 2: Quantitative evaluation of the aggregated disparity map generated by our proposed hybrid method on a subset of Middlebury stereo dataset

Metrics	Cones	Teddy	Baby1	Cloth3
Avg. err	5.96	4.28	3.25	4.46
PSNR (dB)	26.67	28.13	31.87	27.56
Bad 1.0 (%)	25.30	26.10	21.18	27.57
Bad 2.0 (%)	22.04	21.57	16.23	15.49
Run time (s)	73.85	60.02	82.01	87.30

Qualitative results after cost aggregation using guided image filtering are shown in Figure 2, which illustrates enhanced accuracy of the estimated disparity map in Cones, Teddy, Baby1 and Cloth3 from Middlebury dataset images.

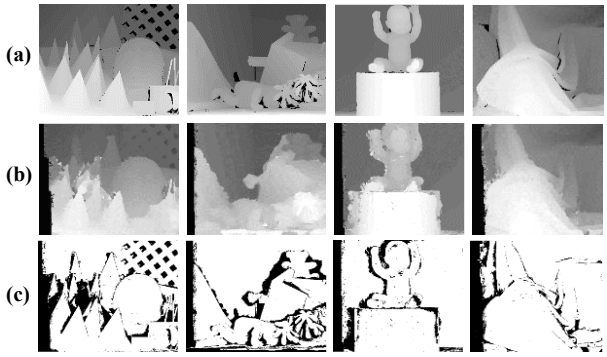


***Figure 2**. Results of the disparity map after cost aggregation on Middlebury dataset images. Cones, Teddy, Baby1, and Cloth3 from left to right. (a) Ground truth, (b) the estimated results from the proposed approach without refinement, and (c) error maps. The occluded regions are included on all the obtained results.*

3.C. Evaluation of the Final Disparity Map on the Middlebury Benchmark

The performance of the final results after postprocessing was evaluated based on the PSNR, Bad1.0, Bad2.0 and Avg. err metrics on the subset of benchmark Middlebury dataset. Table 3 shows quantitative evaluation of the final results and Figure 3 demonstrates the final results for Cones, Teddy, Baby1, and Cloth3 images.

Table 3: Quantitative evaluation of the final disparity map on the subset of the Middlebury stereo images.

Metrics	Cones	Teddy	Baby1	Cloth3
Avg. err	2.36	2.10	1.55	1.97
PSNR (dB)	33.26	32.88	37.84	34.91
Bad 1.0 (%)	18.96	17.72	11.30	19.91
Bad 2.0 (%)	15.35	10.61	6.11	8.38

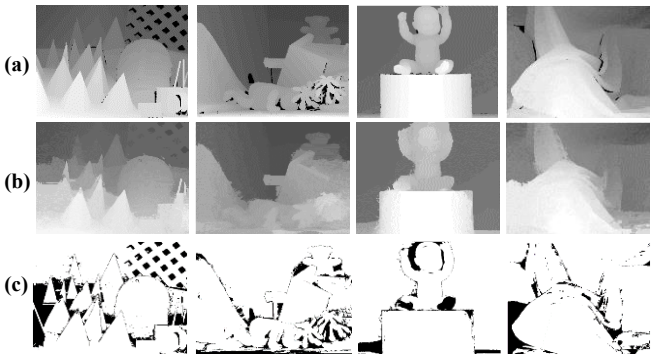


Figure 3. Results of final disparity map on Middlebury dataset images. Cones, Teddy, Baby1, and Cloth3 from left to right. (a) Ground truth, (b) the estimated results from the proposed approach, and (c) error maps. The occluded regions are included on all the obtained results.

We selected 12 stereo image pairs from the Middlebury benchmark dataset to illustrate the performance of the proposed approach on stereo images with a variety of textures. Quantitative metrics are presented in Table 4 and qualitative disparity map evaluations are presented in Figure 4.

Table 4: Quantitative evaluation of the final disparity map on 12 stereo pairs from three different Middlebury dataset.

Images	PSNR (dB)	Percentage of bad pixels (%)		Average error
		Bad01	Bad02	
Cones	33.26	18.96	15.35	2.36
Teddy	32.88	17.72	10.61	2.10
Baby2	31.29	31.77	10.83	2.97
Cloth1	30.03	28.73	15.80	3.57
Cloth3	34.91	19.91	8.38	1.97
Cloth4	29.49	24.69	12.81	3.43
Rocks1	30.70	33.35	12.83	3.02
Rocks2	29.58	25.97	11.07	3.05
Wood1	32.17	26.34	24.67	2.34
Baby1	37.84	11.30	6.11	1.55
Baby3	28.18	18.65	15.49	3.84
Dolls	34.72	36.72	19.35	2.23
AVG	32.06	24.67	13.79	2.71

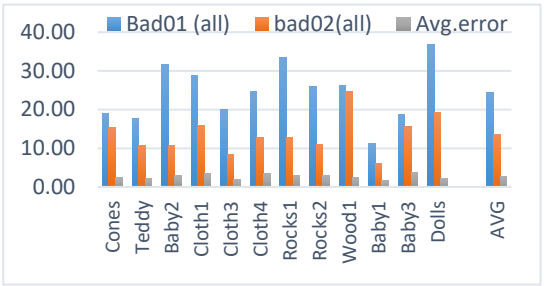


Figure 4. Visualization of the calculated percentage of the bad pixels and the average absolute error of the final disparity map (on 12 pair of images) on the three different Middlebury datasets.

Finally, we evaluated the proposed algorithm based on the current evaluation criteria of the Middlebury dataset (version 3.0). Table 5 shows the quantitative error metrics for the 15 image pairs in the current version 3.0 evaluation dataset.

Table 5: Evaluation of the final disparity map on the current evaluation criteria of the Middlebury data version 3.0.

Images	PSNR (dB)	Percentage of bad pixels (%)		Average error
		Bad01	Bad02	
Adirondack	29.76	33.85	21.68	3.98
ArtL	31.35	46.26	38.65	4.31
Jadelant	25.98	43.26	37.19	27.22
Motorcycle	29.35	37.69	21.66	3.91
MotorcycleE	26.47	42.75	30.07	6.12
Piano	28.15	46.94	35.12	5.23
PianoL	26.61	48.2	38.48	7.44
Pipes	25.74	44.84	38.62	7.52
Playroom	25.76	45.92	34.68	6.16
Playtable	31.17	46.06	28.47	3.57
PlaytableP	31.64	38.5	24.83	3.17
Recycle	31.63	49.54	32.23	3.62
Shelves	27.17	49.49	33.64	5.51
Teddy	32.88	17.72	10.61	2.10
Vintage	28.68	42.87	33.16	10.62
AVG	28.78	42.37	30.72	6.71

3.D. Comparison of the Proposed Approach with Other Algorithms

Performance of our proposed method was compared with that of 12 state-of-the-art non-learning algorithms recently published and reported on Middlebury evaluation version 3.0. Quantitative and qualitative measures of performance of our proposed method and the 12 non-learning methods namely APAP-Stereo[37], MTS[38], SM-AWP[39], FASW[40], MBM[41], IEBIMst[42], MotionStereo[43], DoGGuided[44], SGMEPi[45], IGF[46], ADSM[47], SPS[48] are presented in Table 6. Our proposed method provided the best error metric for 8 of the 15 stereo pairs evaluated.

4. Conclusions

We presented a new hybrid stereo matching algorithm (HCS) that utilized segmentation based *a priori* distribution of disparity estimates and a normalized cross-correlation cost measure. Our framework calculates initial disparity map efficiently by guiding the search locations in the Middlebury benchmark stereo images. We aggregated the obtained cost volume with the edge-preserving guided filtering to improve the accuracy of the results. The results of our proposed method were compared with those of the recent and state-of-the-art approaches. Our proposed hybrid method provided lower disparity errors in 8 of the 15 stereo pairs tested among all the state-of-the-art non-learning methods evaluated. In stereo images with large disparity, our approach provided a relatively higher disparity error. Overall, our approach performed better when compared with the recent state-of-the-art approaches in terms of the calculated average disparity error. In a future work, we will consider global optimization techniques to further enhance the accuracy of the disparity map.

Table 6: Evaluation of the final disparity map on the current evaluation criteria of the Middlebury data version 3.0. Comparison of the final disparity map on the current evaluation set of the Middlebury dataset (version 3.0) with other state-of-the-art algorithms. The best result for each stereo image pair is highlighted in bold.

Approaches	*Adiron*	*ArtL*	*Jadepl*	*Motor*	*MotorE*	*Piano*	*PianoL*	*Pipes*	*Playrm*	*Playt*	*PlaytP*	*Recyc*	*Shelvs*	*Teddy*	*Vintage*	*Average*
Proposed	3.98	**4.31**	**27.2**	3.91	6.12	5.23	7.44	**7.52**	**6.16**	**3.57**	**3.17**	3.62	5.51	**2.10**	10.6	**6.70**
APAP-Stereo[37]	4.90	5.65	44.5	**3.86**	**3.98**	**3.02**	**4.64**	8.02	9.01	4.56	4.04	**3.12**	**3.36**	2.54	6.52	7.45
MTS[38]	19.0	22.5	123	17.5	20.7	13.0	32.0	29.4	26.9	27.4	12.0	17.5	12.1	8.11	27.2	27.2
SM-AWP[39]	10.5	19.9	62.7	11.0	12.5	9.08	29.7	21.1	20.7	9.5	9.75	7.18	11.4	9.44	16.8	17.4
FASW[40]	**2.86**	8.03	34.7	5.44	5.43	5.54	10.8	10.8	7.31	14.5	3.32	2.84	8.70	2.83	**6.79**	8.66
MBM[41]	4.39	8.80	37.6	5.76	5.56	6.67	12.4	11.8	12.9	12.0	6.37	3.67	11.8	3.74	14.1	10.5
IEBIMst[42]	27.3	15.1	55.6	5.54	8.21	6.40	18.9	11.8	18.0	17.9	4.95	5.29	17.1	5.31	10.9	15.2
MotionStereo[43]	5.21	12.0	36.1	8.16	7.92	5.23	7.57	16.9	10.3	12.3	6.19	4.64	9.48	3.95	16.0	10.8
DoGGuided[44]	20.1	28.0	56.5	13.8	16.8	13.4	37.3	23.8	30.3	30.8	13.0	9.13	19.0	13.4	23.6	23.2
SGMEPi[45]	5.65	18.2	30.8	9.18	9.02	8.49	14.7	15.8	21.0	10.7	9.76	5.80	11.0	10.7	31.9	14.1
IGF[46]	4.56	7.33	28.8	5.87	5.91	6.36	12.1	11.5	7.16	27.3	8.64	3.85	9.19	3.30	9.12	10.0
ADSM[47]	14.3	10.6	34.1	6.0	8.0	7.37	20.4	12.1	16.9	25.5	5.84	5.83	17.2	4.11	11.1	13.2
SPS[48]	6.51	15.2	40.0	8.35	8.45	12.0	25.0	16.1	25.2	15.7	12.4	8.81	23.7	8.01	53.7	18.6

Acknowledgement

The first author Ms. Hanieh Shabanian, a Ph.D. student, is currently supported through a graduate assistantship from the Department of Electrical and Computer Engineering at The University of Memphis. The research work was supported in part by the Computational Ocularscience laboratory.

References

[1] C. F. Olson, L. H. Matthies, M. Schoppers, and M. W. Maimone, "Rover navigation using stereo ego-motion," *Rob. Auton. Syst.*, vol. 43, no. 4, pp. 215–229, Jun. 2003, doi: 10.1016/S0921-8890(03)00004-6.

[2] R. D. Eastman and A. M. Waxman, "Using disparity functionals for stereo correspondence and surface reconstruction.," *Comput. Vision, Graph. Image Process.*, vol. 39, no. 1, pp. 73–101, Jul. 1987, doi: 10.1016/S0734-189X(87)80203-7.

[3] N. Uchida, T. Shibahara, T. Aoki, H. Nakajima, and K. Kobayashi, "3D face recognition using passive stereo vision," in *Proceedings - International Conference on Image Processing, ICIP*, 2005, vol. 2, pp. 950–953, doi: 10.1109/ICIP.2005.1530214.

[4] N. Kaempchen, U. Franke, and R. Ott, "Stereo vision based pose estimation of parking lots using 3D vehicle models," 2003, pp. 459–464, doi: 10.1109/ivs.2002.1187993.

[5] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1–3, pp. 7–42, 2002, doi: 10.1023/A:1014573219977.

[6] H. Ishikawa and D. Geiger, "Occlusions, discontinuities, and epipolar lines in stereo," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1998, vol. 1406, pp. 232–248, doi: 10.1007/BFb0055670.

[7] J. Sun, N. N. Zheng, and H. Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 787–800, Jul. 2003, doi: 10.1109/TPAMI.2003.1206509.

[8] K. Ambrosch, M. Humenberger, W. Kubinger, and A. Steininger, "SAD-Based Stereo Matching Using FPGAs," in *Embedded Computer Vision*, Springer London, 2008, pp. 121–138.

[9] L. De-Maeztu, A. Villanueva, and R. Cabeza, "Stereo matching using gradient similarity and locally adaptive support-weight," *Pattern Recognit. Lett.*, vol. 32, no. 13, pp. 1643–1651, Oct. 2011, doi: 10.1016/j.patrec.2011.06.027.

[10] M. Marghany and M. Hashim, "3D stereo reconstruction using sum square of difference matching algorithm," *Sci. Res. Essays*, vol. 6, no. 30, pp. 6404–6423, 2011, doi: 10.5897/SRE11.1661.

[11] R. Zabih and J. Woodfill, "Non-parametric local transforms for computing visual correspondence," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1994, vol. 801 LNCS, pp. 151–158, doi: 10.1007/bfb0028345.

[12] Y. S. Heo, K. M. Lee, and S. U. Lee, "Robust Stereo matching using adaptive normalized cross-correlation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 807–822, 2011, doi: 10.1109/TPAMI.2010.136.

[13] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, 2013, doi: 10.1109/TPAMI.2012.213.

[14] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 2, pp. 504–511, 2013, doi: 10.1109/TPAMI.2012.156.

[15] G. S. Hong, J. K. Park, and B. G. Kim, "Near real-time local stereo matching algorithm based on fast guided image filtering," in *Proceedings of the 2016 6th European Workshop on Visual Information Processing, EUVIP 2016*, 2016, doi: 10.1109/EUVIP.2016.7764595.

[16] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 120–129, Jan. 2015, doi: 10.1109/TIP.2014.2371234.

[17] G. S. Hong, M. S. Koo, A. Saha, and B. G. Kim, "Efficient local stereo matching technique using weighted guided image filtering (WGIF)," in *2016 IEEE International Conference on Consumer Electronics, ICCE 2016*, 2016, pp. 484–485, doi: 10.1109/ICCE.2016.7430699.

[18] S. Zhu and L. Yan, “Local stereo matching algorithm with efficient matching cost and adaptive guided image filter,” *Vis. Comput.*, vol. 33, pp. 1087–1102, 2017, doi: 10.1007/s00371-016-1264-6.

[19] K. J. Yoon and I. S. Kweon, “Adaptive support-weight approach for correspondence search,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 650–656, Apr. 2006, doi: 10.1109/TPAMI.2006.70.

[20] Q. Yang, P. Ji, D. Li, S. Yao, and M. Zhang, “Fast stereo matching using adaptive guided filtering,” *Image Vis. Comput.*, vol. 32, no. 3, pp. 202–211, Mar. 2014, doi: 10.1016/j.imavis.2014.01.001.

[21] G.-S. Hong and B.-G. Kim, “A local stereo matching algorithm based on weighted guided image filtering for improving the generation of depth range images,” 2017, doi: 10.1016/j.displa.2017.07.006.

[22] C. Zhu and Y.-Z. Chang, “Efficient Stereo Matching Based on Pervasive Guided Image Filtering,” 2019, doi: 10.1155/2019/3128172.

[23] C. Han, W. Liu, L. Jin, S. Jiang, and H. Li, “Adaptive Weight Based Sparse Block Aggregation Algorithm for Stereo Matching,” in *9th International Conference on Information Science and Technology, ICIST 2019*, 2019, pp. 408–412, doi: 10.1109/ICIST.2019.8836748.

[24] A. Oppenheim and R. S. Processing, “From frequency to quefrency: A history of the cepstrum,” *ieeexplore.ieee.org*.

[25] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. 2004.

[26] R. I. Hartley, “Theory and practice of projective rectification,” *Int. J. Comput. Vis.*, vol. 35, no. 2, pp. 115–127, Nov. 1999, doi: 10.1023/A:1008115206617.

[27] A. K. Jain and F. Farrokhnia, “Unsupervised texture segmentation using Gabor filters,” *Pattern Recognit.*, vol. 24, no. 12, pp. 1167–1186, Jan. 1991, doi: 10.1016/0031-3203(91)90143-S.

[28] J. Shi and C. Tomasi, “Good features to track,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600, doi: 10.1109/cvpr.1994.323794.

[29] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-Up Robust Features (SURF),” *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, Jun. 2008, doi: 10.1016/j.cviu.2007.09.014.

[30] C. Harris and M. Stephens, “A Combined Corner and Edge Detector,” 2013, pp. 23.1-23.6, doi: 10.5244/c.2.23.

[31] D. G. Viswanathan, “Features from Accelerated Segment Test (FAST) Deepak Geetha Viswanathan 1.,” 2011.

[32] M. Balasubramanian *et al.*, “A framework for detecting glaucomatous progression in the optic nerve head of an eye using proper orthogonal decomposition,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 781–793, 2009, doi: 10.1109/TITB.2009.2020158.

[33] D. Scharstein and R. Szeliski, “High-accuracy stereo depth maps using structured light,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003, vol. 1, doi: 10.1109/cvpr.2003.1211354.

[34] H. Hirschmuller, D. S.-2007 I. C. on, and undefined 2007, “Evaluation of cost functions for stereo matching,” *ieeexplore.ieee.org*.

[35] D. Scharstein and C. Pal, “Learning conditional random fields for stereo,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007, doi: 10.1109/CVPR.2007.383191.

[36] “vision.middlebury.edu/stereo.” [Online]. Available: http://vision.middlebury.edu/stereo/. [Accessed: 03-Mar-2020].

[37] M. G. Park and K. J. Yoon, “As-planar-as-possible depth map estimation,” *Comput. Vis. Image Underst.*, vol. 181, pp. 50–59, Apr. 2019, doi: 10.1016/j.cviu.2019.02.001.

[38] R. Brandt, N. Strisciuglio, N. P.-P. recognition letters, and undefined 2020, “Efficient binocular stereo correspondence matching with 1-D Max-Trees,” *Elsevier*.

[39] S. S. A. Razak, M. A. Othman, and A. F. Kadmin, “The effect of adaptive weighted bilateral filter on stereo matching algorithm,” *Int. J. Eng. Adv. Technol.*, vol. 8, no. 3, pp. 284–287, 2019.

[40] W. Wu, H. Zhu, S. Yu, and J. Shi, “Stereo Matching with Fusing Adaptive Support Weights,” *IEEE Access*, vol. 7, pp. 61960–61974, 2019, doi: 10.1109/ACCESS.2019.2916035.

[41] Q. Chang and T. Maruyama, “Real-Time Stereo Vision System: A Multi-Block Matching on GPU,” *IEEE Access*, vol. 6, pp. 42030–42046, Jul. 2018, doi: 10.1109/ACCESS.2018.2859445.

[42] C. He, C. Zhang, Z. Chen, and S. Jiang, “Minimum spanning tree based stereo matching using image edge and brightness information,” in *Proceedings - 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2017*, 2018, vol. 2018-January, pp. 1–5, doi: 10.1109/CISP-BMEI.2017.8302008.

[43] J. Valentin *et al.*, “Depth from motion for smartphone Ar,” in *SIGGRAPH Asia 2018 Technical Papers, SIGGRAPH Asia 2018*, 2018, vol. 37, no. 6, pp. 1–19, doi: 10.1145/3272127.3275041.

[44] M. Kitagawa, I. Shimizu, and R. Sara, “High accuracy local stereo matching using DoG scale map,” in *Proceedings of the 15th IAPR International Conference on Machine Vision Applications, MVA 2017*, 2017, pp. 258–261, doi: 10.23919/MVA.2017.7986850.

[45] D. Scharstein, T. Taniai, and S. N. Sinha, “Semi-global stereo matching with surface orientation priors,” in *Proceedings - 2017 International Conference on 3D Vision, 3DV 2017*, 2018, pp. 215–224, doi: 10.1109/3DV.2017.00033.

[46] R. A. Hamzah, H. Ibrahim, and A. H. Abu Hassan, “Stereo matching algorithm based on per pixel difference adjustment, iterative guided filter and graph segmentation,” *J. Vis. Commun. Image Represent.*, vol. 42, pp. 145–160, Jan. 2017, doi: 10.1016/j.jvcir.2016.11.016.

[47] N. Ma, Y. Men, C. Men, and X. Li, “Accurate Dense Stereo Matching Based on Image Segmentation Using an Adaptive Multi-Cost Approach,” *Symmetry (Basel).*, vol. 8, no. 12, p. 159, Dec. 2016, doi: 10.3390/sym8120159.

[48] C. LeGendre, K. Batsos, and P. Mordohai, “High-Resolution Stereo Matching based on Sampled Photoconsistency Computation,” *Br. Mach. Vis. Conf.*, 2017.

Author Biography

Hanieh Shabanian received her M.S. in Computer Engineering from the Department of Electrical and Computer Engineering (EECE) at the University of Memphis (UofM). Since then she has worked as a research assistant in the Computational Ocularscience Laboratoryand pursuing her Ph.D. degree in the EECE at U of M. Her research interest and dissertation projects are at the intersection of ocular imaging, computer vision, image processing and machine learning.

Madhusudhanan Balasubramanian, Ph.D. is the Director and PI of the Computational Ocularscience laboratory, and serves as Associate Professor in the EECE department at U of M. His current research interests include computer vision, glaucoma, retinal physiology, ocular drug delivery, biophysical and multi-physics modeling.

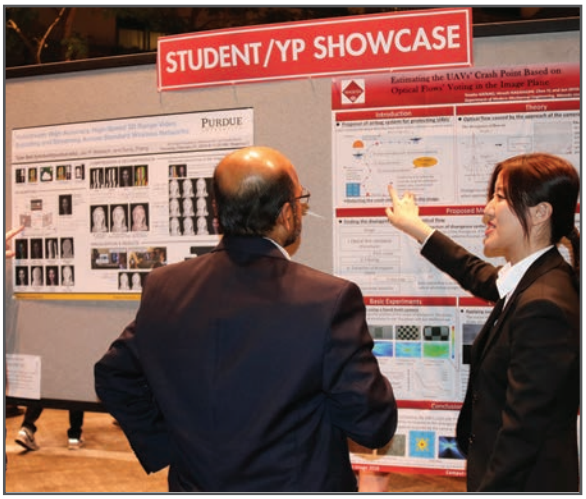
JOIN US AT THE NEXT EI!

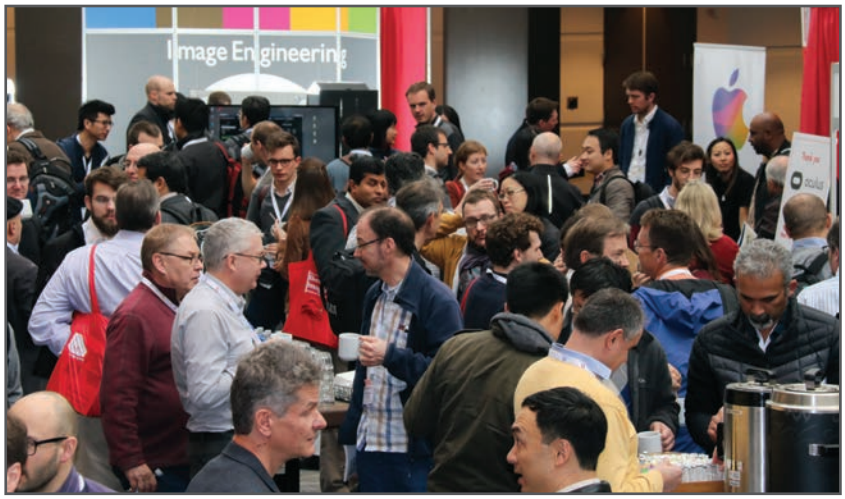
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org