01-02

16th International Conference on Machine Vision Applications (MVA)
National Olympics Memorial Youth Center, Tokyo, Japan, May 27-31, 2019.

Reconstruction with Guided PatchMatch Stereo

Trevor Gee
University of Auckland
Department of Computer Science

Patrice Delmas
University of Auckland
Department of Computer Science

Abstract

Stereo matching is an operation that calculates a dense set of correspondences from a pair of images engineered to conform to canonical epipolar geometry. Stereo matching is a key component in many 3D reconstruction systems, thus it is a heavily studied area in the field of computer vision. Stereo matching, in the general case, is a difficult problem that remains unsolved to a satisfactory level for most scientists in the field. However, due to the large amount of attention given to this problem, there exists many algorithms, each with their own advantages and problems. This work explores the notion of combining the outputs of two well known stereo matching algorithms (symmetric dynamic programming stereo and patch match stereo) using a statistical framework, with the hope that the strategy will preserve the advantages of both algorithms while overcoming some of the weaknesses. Experiments were performed both on images from the Middlebury data set and ones captured for this study using a stereo camera. The results indicate that the proposed strategy is promising.

1 Introduction

Stereo matching, for 3D reconstruction, has been an important problem in the Computer Vision community dating back to 1976 [1]. While several solutions perform extremely well in limited well-defined contexts, it is still considered to be an unsolved problem in the general case. Websites like Middlebury Stereo [2] and KITTI [3] are just some of the resources available dedicated to finding a solution for the general stereo matching problem. Few other problems in the field of computer vision have received the same attention.

Over the years, a huge variety of approaches have been proposed from simple block matching [4] to sophisticated Markov Random Field (MRF) formulations [5]. Modern stereo matching algorithms are now incorporating the latest advances in machine learning [6]. Key problems in stereo matching include the lighting artifacts (shadows for example), the reflectance properties of materials (transparent for example), occlusions (essentially monocular pixels in the scene) and correctly interpreting homogeneous surfaces (many different surface profiles appear "correct" as per the image data).

A feature of the collective stereo algorithms, is that they tend to have different strengths and weaknesses. This leads to the notion of *guided stereo*, a strategy that attempts to combine the strengths of two stereo matching algorithms by improving the output of a first algorithm with a second algorithm. This work builds on work by Nguyen et al [7], who used traditional Sum of Absolute Differences (SAD) block matching to guide a Symmetric Dynamic Stereo algorithm [8] (SDPS). Guidance was achieved by adding a bias to the data term of the second algorithm, causing it to favor the results of the first algorithm unless it found a better result.

The novelty of this work is that it proposes a new strategy of guidance, based on statistical distributions of disparities in the local neighborhood, where it is postulated that the proposed algorithm is less likely to by influenced by outliers than the approach of Nguyen et al. Also in this work, it was decided to focus on blending Symmetric Dynamic Programming Stereo (SDPS) [8] and PatchMatch Stereo [9]. The reason that these two algorithms were chosen, is that these two algorithms have complementary advantages: SDPS has a global optimization that PatchMatch lacks, while PatchMatch eliminates the tell-tale streaking artifacts associated with the 1-D nature of SDPS and gives SDPS sub-pixel disparity approximation.

In the next Section 2, a literature review is presented. In Section 3, the reconstruction pipe-line and the guided patch-match algorithm is proposed. In Section 4, the experimental methodology is outlined. Finally experimental results are presented in Section 5 and conclusion are in Section 6.

2 Literature Review

Stereo matching finds a dense set of corresponding pixels in the common field of view between two images. If both images have been rectified to match canonical epipolar geometry, then corresponding pixels are always on the same row. One advantage of having correspondences on the same row, is that correspondences may be stored as an offset map, known as a disparity map.

A typical strategy for finding correspondences is to find pixels with a similar color or intensity. Unfortunately, a single pixel based search is typically too basic for stereo matching, therefore a matching cost function typically relies on matching small local windows across left and righ images. Typical match costs can be based on the Sum of Absolute Differences (SAD) [10], Sum of Squared Differences [11] and Normalised Cross Correlation (NCC) [12] to name a few.

Illumination differences between left and right stereo images, however, may still be a problem, and cost functions relying on mutual information [13], the census transform [14] and image gradients [15] have shown some invariance to illumination differences between stereo pairs.

Window-based matching also tends to suffer from fattening artifacts, since a window searching typically falsely assumes that all pixels in the sliding window have the same disparity. One strategy for overcoming this is to use a color weight to filter out elements that have a different color with respect to the central point of interest. This strategy is known as Adaptive Support Weights [16].

Stereo matching algorithms can operate fully as a correspondence search based on local search windows, however, these algorithms struggle with texture-less regions with large ambiguity in matching and occlusions (pixels that are only visible in one of the stereo images, typically found close to edges). A typical approach in dealing with this uncertainty is to impose the heuristic that texture-less regions and monocular pixels are smooth, that is to say, that disparity values are nomrally similar to their neighbors. The smoothness constraint is typically enforced with a smoothing term added to the cost function. As smoothness often needs to be propagated over some distance within images, these smoothing costs tend to be globally optimized. Unfortunately, in 2-D, this becomes an NP-hard problem. It is solvable in 1-D, and algorithms such as SDPS [8] and 1-D Belief Propagation [17] take advantage of this.

There are several algorithms that attempt to approximate a full 2-D global optimized stereo matching solution. These include Graph-Cuts Stereo [5], Semi-Global Matching [18] and Belief Propagation [19]. Unfortunately, most of these algorithms are quite slow and thus researchers have been looking for algorithms that enforce global constraints in a quicker way. Efficient large-scale stereo (ELAS) [20] does this by fitting planes to a sparse network of reliable correspondences and use this to guide a local search for correspondences. Another popular approach is to use color segmentation [21].

The current top stereo matching algorithms, according to the Middlebury stereo website [2], are using machine learning to encode many of the complexities in stereo matching problems. A convolutional neural network based image patch matching of Zbontar et al. [6] is an example of such work.

This work moves away from the traditional path of stereo matching algorithm research, in that, instead of proposing a completely new technique, it looks at what already exists and attempts to use the knowledge acquired from previous research to engineer a new superior algorithm. As mentioned in the introduction, this work attempts to continue on work published by Nguyen et al [7], which investigates the notion of guided stereo matching.

3 Proposed Algorithm

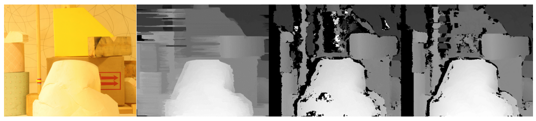

Figure 1. **Guided Patchmatch Stereo:** (Left) A color frame from the Middlebury dataset. (Middle Left) A disparity map generated from SDPS. (Middle Right) The result from PatchMatch Stereo. (Right) The Guided Patch match result.

Given a left rectified image I_r and a right rectified image I'_r, along with a proposed solution disparity map D_{guide}, the goal of the proposed algorithm is to find disparity map D, which is potentially an improved version of D_{guide}.

3.0.1 Parameterisation

Before refining the disparity map D_{guide}, it is necessary to define how to parametrise the problem. The authors of Patch Match Stereo parameterised their solutions space directly from the 3-D plane definition $d_i = a_i u + b_i v + c_i$ where (a_i, b_i, c_i) are the plane gradient parameters for pixel i, $\mathbf{p} = [u, v]^T$ represents the 2-D coordinate within the plane and d_i is the disparity value assigned to pixel i. This plane equation encodes the position and tilt of the search window, allowing it to be positioned and warped accordingly during a search for better parameters. In this proposed implementation, it was decided to parameterise the problem more directly for the sake of simplicity. The proposed parameterisation takes the form of the parameter vector $\mathbf{S_i} = [d_i, \theta_i, \phi_i]$, where d_i is the disparity value of pixel i and θ_I and ϕ_i are the angle tilts in the X- and Y-directions of the local search window applied as follows:

Given local search window centered on pixel $\mathbf{p_i} = [u_i, v_i]^T$ with offsets j and k, then the value at $I(u_i + j, v_i + k)$ is compared with the value in the corresponding window $I'(u_i - d_i + jcos(\theta_i), v_i + kcos(\phi_i))$. Bilinear interpolation is used to deal with non-integer values. It was found that both parameterisations produced similar results.

3.0.2 Cost Function

The cost function here generally follows the one proposed for PatchMatch Stereo as:

$$E_{cost} = C(\mathbf{p_i}, \mathbf{S_i}, I, I') = \sum_{j=-b}^{b} \sum_{k=-b}^{b} W(\mathbf{p_i}, j, k, I)\rho(\mathbf{p_i}, j, k, \mathbf{S_i}, I, I') \tag{1}$$

where the function W determines the support weights as:

$$W(\mathbf{p_i}, j, k, I) = e^{-\frac{\|I(u_i, v_i) - I(u_i + j, v_i + k)\|}{\gamma}} \tag{2}$$

and ρ determines the actual differences per pixel as:

$$\rho(\mathbf{p_i}, j, k, \mathbf{S_i}, I, I') = (1 - \alpha) \cdot \min(\|I_p - I'_q\|, \tau_{col}) + \alpha \cdot \min(\|\nabla I_p - \nabla I'_q\|, \tau_{grad}) \tag{3}$$

where α is a "blend" parameter $\{0 \leqslant \alpha \leqslant 1\}$, $I_p = I(u_i + j, v_i + k)$ and $I_q = I'(u_i - d + jcos(\theta_1), v_i + kcos(\phi_2))$, $\nabla I_p = \nabla I(u_i + j, v_i + k)$ where ∇I is the gradient image of I and $\nabla I'_q = \nabla I'(u_i - d_i + jcos(\theta_i), v_i + kcos(\phi_i))$, $\nabla I_p = \nabla I(u_i + j, v_i + k)$ where $\nabla I'$ is the gradient image of I'. The gradient images where produced by combining X-gradient images and Y-gradient images generated by the Sobel [22] operator.

3.0.3 Guided Cost Function

The cost function C, defined above, provides a means to determine the fit parameters $\mathbf{S_i}$ have with the corresponding pair at point $\mathbf{p_i}$ in I and $[u_i - d, v_i]^T$ in I'. If the value of C is small then the probability of a good match of parameters is high and vice-versa. The probability of a good match $Pr(M)$, given C (assuming a Gaussian Distribution) is proportional to:

$$Pr(M|C) \propto e^{-\frac{\|C\|}{\gamma}} \tag{4}$$

where γ is a guess at the variance of the disparity in the local window.

Another aspect considered is the information that disparity map D_{guide} gives about the likely disparity d_i for pixel $\mathbf{p_i}$. Patch Match Stereo makes the assumption that disparity values in a small local region are likely to be similar. Given the assumption that noise in disparity SDPS disparity maps can be modelled as a Gaussian distribution, then for location $\mathbf{p_i}$ in D_{guide}, given the radius r, it is relatively straightforward to determine the mean μ_r and standard deviations σ_r of the disparity values that surround $\mathbf{p_i}$. The probability of a good match $Pr(M|r)$ given the radius r should be proportional to:

$$Pr(M|r) \propto e^{-\frac{\|di - \mu_r\|}{2\sigma^2}} \tag{5}$$

The match probability of Eqn. 4 and the distribution probably of Eqn. 5 can be combined into a single cost function as follows:

$$E(\mathbf{p_i}, \mathbf{S_i}, I, I') = \beta C(\mathbf{p_i}, \mathbf{S_i}, I, I') - \log\left[\exp\left(\frac{-\|d_i - \mu_r\|}{2\sigma_r^2}\right)\right] \tag{6}$$

where β controls the weighting difference between the first and second terms. It was decided to call this cost function the *Guided Cost Function*, since it takes into account the match score, but is also guided by the disparity distributions in the guiding disparity map. It should be noted that some inspiration for this formulation comes work published by Geiger et al [20].

3.0.4 Refinement

The Guided Patchmatch algorithm builds two parameter maps, S for I and S' for I'. Each pixel $\mathbf{p_i}$ corresponds to the triplet $\mathbf{S_i} = [d_i, \theta_i, \phi_i]$. Each $\mathbf{S_i}$ in S is initialised as $\mathbf{S_i} = [D_{guide}(\mathbf{p_i}), 0, 0]$ where D_{guide} is the guiding disparity map. Each $\mathbf{S'_i}$ in S' is initialised as $\mathbf{S_i} = [D_{guide}([u_i - D_{guide}(\mathbf{p_i}, v_i)]^T), 0, 0]$. Optimisation of the parameters follows the normal Patchmatch Stereo operations of Refinement, Spatial Propogation and View Propagation. The reader is refered to the PathMatch Stereo paper [9] for more details on these operations. It should be noted however that during the operation the following range limits are imposed:

- $\{di \mid d_i - 3\sigma_r \leq di \leq d_i + 3\sigma_r\}$, since disparity values are expected to be within 3 standard deviations of the local distribution by the assumed Gaussian distribution.
- $\{\theta_i \mid \frac{-\pi}{2} < \theta_i < \frac{\pi}{2}\}$ to avoid the search window flipping.
- $\{\phi_i \mid \frac{-\pi}{2} < \phi_i < \frac{\pi}{2}\}$ to avoid the search window flipping.

4 Experimental Methodology

4.1 Algorithms

A basis of comparison is required and therefore the following state-of-the-art stereo matching algorithms were used as a benchmark.

- **Efficient Large-Scale Stereo** [20] (ELAS): The implementation provided by the author was used.
- **Block Matching Stereo** [23] (BM): The highly efficient OpenCV version was used.
- **Semi-global Block Matching** [18] (SGBM): OpenCV's version was used here too.
- **Symmetrical Dynamic Programming** [8] (SDPS): The IVS Lab version was used.
- **Patch Match Stereo** [9] (PMS): The IVS lab's CUDA version was used.

Figure 3. **Stereo Camera:** The FujiFilm FinePix REAL 3D W3

4.2 Middlebury Experiments

Experiments were performed against a variety of stereo pairs chosen from the Middlebury [2]. The goal of these experiments was to determine how well the proposed algorithm performs against other state-of-the-art algorithms. The images of the Middlebury dataset tend to be rather idealistic, with perfect rectification and illumination, therefore it provides a perfect test basis to assess optimal performance. The Middlebury database also contains "ground truth" images for evaluation. Accuracy was assessed by directly finding the variance between the disparity values of the ground truth and that of the proposed algorithm. In cases where the tested stereo matching algorithm was unable to compute a match, values were excluded from the average error. A secondary coverage score is provided to indicate the percentage of pixels skipped by the algorithm.

4.3 Orbbec Astra Experiments

Figure 2. **Astra Orbbec:** An infra-red active depth sensor camera

In order to assess the performance of the algorithm in indoor scenes, experiments were performed using a Orbbec Astra camera (Fig. 2). The Orbbec Astra has an RGB camera and an infra-red based active depth sensor and projector which are produce depth maps of indoor scenes. Our tests determined the depth sensor accuracy to be ±3 mm at a range of 1 m.

The idea behind the Orbbec Astra experiments was to compare the 3-D reconstructions produced from the disparity maps of the proposed algorithm to the reconstructions produced by the Orbbec Astra (ground-truth).

The level of consistency between the point cloud of the Orbbec Astra and the proposed algorithm was measured using the Hausdorff distance metric [24].

4.4 Outdoor Experiments

In order to assess performance in real outdoor scenes, a database was acquired using a calibrated W3 camera (see Fig. 3). Images were processed by a 3-D pipeline using the proposed algorithm.

5 Results

5.1 Middlebury

Table 1, provides the results of Efficient Large-Scale Stereo [20] (ELAS), Block Matching Stereo [23] (BM), Semi-global Block Matching [18] (SGBM), Patch Match Stereo [9] (PMS), Symmetrical Dynamic Programming [8] (SDPS) and the proposed algorithm Guided Block Patching Stereo (GPMS) to stereo pairs from Middlebury. Thumbnails of the images used are shown in Fig. 4. The image numbers (the first column "No" of Table 1) corresponds with Fig. 4, with image 0 at the top-left, and the remaining images ordered column-by-column and then row-by-row there after. Each stereo matching algorithm has an associated column "Err"indicating the average disparity error from the ground truth in pixels. The various algorithms also marked pixels as "not found" and these were excluded from the average disparity error. The column marked "Cov" indicates percentage coverage as the ratio between the pixels that were found to have valid disparity values with respect to the total pixel count of the image.

Figure 4. **Middlebury Database:** Thumbnail images of the left images of the stereo pairs from Middlebury database used in this work.

From Table 1 it is clear the ELAS performs the best on the Middlebury images. However, the proposed algorithm is a close second and tends to do better on the coverage metric on average. It is also notable that GPMS does generally outperform SDPS and PMS which are the two algorithms that it is composed of.

5.2 Orbbec Astra

Experiments were conducted with four scenes with the Orbbec Astra to capture images and ground truth depth map. The images featured a pillow covered by a tracksuit top, a shoe, a vacuum cleaner and bedroom (see Fig. 4). The vacuum cleaning and the bedroom scenes were deemed particular difficult due to the presence of homogeneous surfaces. The top algorithms from the Middlebury experiments were chosen for these experiments (Efficient Large-Scale Stereo [20] (ELAS), Semi-global Block Matching [18] (SGBM) and the proposed algorithm Guided Block Patching Stereo (GPMS)).

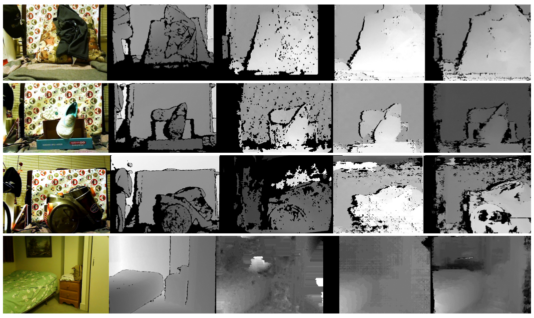

Figure 5. **Orbbec Astra Images:** (Left) The color frame from the Orbbec Astra. (Middle Left) The depth map from the Orbbec Astra. (Middle Middle) The ELAS disparity map. (Middle Right) The SGBM disparity map. (Right) The GPMS disparity map. The scale difference between the disparity map and the depth map is due to the scaling after calibration.

Table 1. Disparity Errors (Pixels) and Coverage (%) of Stereo Matching Algorithms applied to Middlebury Stereo Pairs

	ELAS		BM		SGBM		PMS		SDPS		GPMS	
No	**Err**	**Cov**	**Err**	**Cov**	**Err**	**Cov**	**Err**	**Cov**	**Err**	**Cov**	**Err**	**Cov**
0	2.1	0.77	2.6	0.62	3.8	0.72	3.0	0.83	4.7	0.97	2.9	0.81
1	3.0	0.79	3.1	0.60	4.2	0.73	3.8	0.84	4.2	0.98	3.4	0.83
2	1.3	0.77	3.2	0.57	4.8	0.66	3.0	0.80	3.2	0.97	2.2	0.74
3	3.6	0.66	6.7	0.53	7.8	0.63	6.0	0.67	9.7	1.00	6.3	0.51
4	1.2	0.83	1.1	0.67	2.9	0.74	1.2	0.88	2.6	1.00	1.3	0.80
5	1.3	0.76	1.1	0.60	2.1	0.67	1.8	0.78	11.4	0.94	2.7	0.65
6	1.6	0.78	2.2	0.60	3.0	0.69	2.5	0.83	7.2	0.99	2.0	0.70
7	3.0	0.69	3.7	0.45	6.8	0.51	6.8	0.71	3.3	0.98	2.3	0.63
8	0.9	0.80	0.8	0.64	2.3	0.71	1.2	0.84	4.0	0.99	1.1	0.57
9	1.5	0.81	1.8	0.62	2.7	0.72	2.2	0.84	3.0	0.99	2.0	0.73
10	4.1	0.73	4.6	0.49	11.7	0.56	6.3	0.71	8.0	0.97	4.8	0.61
11	2.7	0.64	6.2	0.51	4.4	0.73	4.2	0.74	7.8	1.00	4.1	0.71
12	2.0	0.79	2.0	0.61	3.9	0.71	3.3	0.81	4.1	0.99	3.5	0.72
13	2.5	0.75	5.2	0.49	6.1	0.61	4.4	0.78	6.2	0.99	3.7	0.61
14	1.0	0.81	1.2	0.67	1.6	0.73	1.4	0.84	3.3	0.95	1.4	0.78
15	1.3	0.80	1.8	0.61	6.1	0.62	1.8	0.82	14.0	0.98	3.5	0.85

Table 2. Hausdorff distance (in millimeters) between Orbbec Astra point clouds and those generated from Reconstruction Pipeline

Image	ELAS	SGBM	GPMS
1	7.1	10.67	6.9
2	6.2	3.9	4.6
3	31.4	25.9	25.9
4	110.1	75.5	48.3

Table 2 shows the average Hausdorff distance (in mm) found between the point cloud generated between the Orbbec Astra ground truth and models made using the proposed pipeline with various stereo matching algorithms. The GPMS algorithm frequently produced the most accurate results. The distance between the scene and the camera was approximately 1 meter.

5.3 W3 Experiments

The final set of experiments performed used the proposed stereo reconstruction pipeline along with GPMS to generate models. Some of these models are shown in Fig. 6

Figure 6. **Reconstructed scenes:** (Left) Original left frame on stereo pair. (Right) The 3-D reconstruction.

6 Conclusions

In this work a Guided Patch Match Stereo (GPMS) algorithm was proposed. The algorithm initially generated a base disparity map using the SDPS algorithm. This result was then refined using local neighborhood statistic from the base disparity map and a random search proposed by the patchmap stereo algorithm to refine the results.

The resultant disparity map was a noticeable improvement on the SDPS algorithm, resultant in much smoother looking disparity maps without the streak artifacts caused by the one dimensional behavior of the SDPS algorithm. The proposed algorithm was a regular best performing algorithm on both the Middlebury images and the indoor experiments with the Orbbec Astra. The algorithm was also used to produce reasonably good models from binocular stereo images from outdoor scenes.

The algorithm is still limited, however from typical issues that plague stereo matching algorithms, such as the inability to perform well on featureless surfaces. However since the algorithm is based upon well known algorithms, this work really represents a technique that may be use to engineer good stereo matching solutions from existing algorithms by changing together algorithms that have desirable characteristics for the problems at hand.

In future work, more comprehensive testing needs to be done and experimentation with other combinations of algorithms needs to be tried to prove the scalability of the approach.

References

[1] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 194, no. 4262, pp. 283–287, 1976.

[2] D. Scharstein and R. Szeliski, "Middlebury stereo vision page," 2002.

[3] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[4] Y.-S. Chen, Y.-P. Hung, and C.-S. Fuh, "Fast block matching algorithm based on the winner-update strategy," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1212–1222, 2001.

[5] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 2, pp. 508–515, IEEE, 2001.

[6] J. Zbontar and Y. LeCun, "Stereo matching by training a convolutional neural network to compare image patches," *Journal of Machine Learning Research*, vol. 17, no. 1-32, p. 2, 2016.

[7] M. Nguyen, Y. H. Chan, P. Delmas, and G. Gimel'farb, "Symmetric dynamic programming stereo using block matching guidance," in *Image and Vision Computing New Zealand (IVCNZ), 2013 28th International Conference of*, pp. 88–93, IEEE, 2013.

[8] G. Gimel'farb, "Probabilistic regularisation and symmetry in binocular dynamic programming stereo," *Pattern Recognition Letters*, vol. 23, no. 4, pp. 431–442, 2002.

[9] M. Bleyer, C. Rhemann, and C. Rother, "Patchmatch stereo-stereo matching with slanted support windows.," in *Bmvc*, vol. 11, pp. 1–11, 2011.

[10] K. Ambrosch, M. Humenberger, W. Kubinger, and A. Steininger, "Sad-based stereo matching using fpgas," in *Embedded Computer Vision*, pp. 121–138, Springer, 2009.

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

[12] Y. S. Heo, K. M. Lee, and S. U. Lee, "Robust stereo matching using adaptive normalized cross-correlation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 807–822, 2011.

[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.

[14] W. S. Fife and J. K. Archibald, "Improved census transforms for resource-optimized stereo vision," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 1, pp. 60–73, 2013.

[15] G.-Q. Wei, W. Brauer, and G. Hirzinger, "Intensity- and gradient-based stereo matching using hierarchical gaussian basis functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1143–1160, 1998.

[16] K.-J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 650–656, 2006.

[17] R. Gong, *Belief propagation based stereo matching with due account of visibility conditions.* PhD thesis, University of Auckland, 2011.

[18] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, pp. 807–814, IEEE, 2005.

[19] J. Sun, N.-N. Zheng, and H.-Y. Shum, "Stereo matching using belief propagation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 25, no. 7, pp. 787–800, 2003.

[20] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Asian conference on computer vision*, pp. 25–38, Springer, 2010.

[21] M. Bleyer and M. Gelautz, "A layered stereo matching algorithm using image segmentation and global visibility constraints," *ISPRS Journal of Photogrammetry and remote sensing*, vol. 59, no. 3, pp. 128–150, 2005.

[22] I. Sobel, "An isotropic 3× 3 image gradient operator," *Machine vision for three-dimensional scenes*, pp. 376–379, 1990.

[23] S. Dröppelmann, M. Hueting, S. Latour, and M. van der Veen, "Stereo vision using the opencv library," 2010.

[24] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317. Springer Science & Business Media, 2009.