

THE EFFECT OF SHADOW AREA ON SGM ALGORITHM AND DISPARITY MAP REFINEMENT FROM HIGH RESOLUTION SATELLITE STEREO IMAGES

Nurollah Tatar ^{a*}, Mohammad Saadatesresht^a, Hossein Arefi^a

^aSchool of Surveying and Geospatial Engineering, College of Engineering, University of Tehran
 (n.tatar, msaadat, hossein.arefi @ut.ac.ir)

KEY WORDS: Shadow, Semi Global Matching, High Resolution Satellite Stereo Images, Disparity Map Refinement.

ABSTRACT:

Semi Global Matching (SGM) algorithm is known as a high performance and reliable stereo matching algorithm in photogrammetry community. However, there are some challenges using this algorithm especially for high resolution satellite stereo images over urban areas and images with shadow areas. As it can be seen, unfortunately the SGM algorithm computes highly noisy disparity values for shadow areas around the tall neighborhood buildings due to mismatching in these lower entropy areas. In this paper, a new method is developed to refine the disparity map in shadow areas. The method is based on the integration of potential of panchromatic and multispectral image data to detect shadow areas in object level. In addition, a RANSAC plane fitting and morphological filtering are employed to refine the disparity map. The results on a stereo pair of GeoEye-1 captured over Qom city in Iran, shows a significant increase in the rate of matched pixels compared to standard SGM algorithm.

1. INTRODUCTION

Dense stereo matching is the primary step in generating digital surface model from satellite images. The methods are categorized as Local (area and feature based) algorithms, Global and Semi-Global matching algorithms (Hirschmüller, 2008; Scharstein and Szeliski, 2002). The weakness of Local matching algorithms is in occlusion boundaries due to the depth discontinuity. Global matching algorithms are formulated as minimizing energy function, which includes discontinuity, smoothness, and occlusion (Brown et al., 2003). These algorithms usually have high computation cost. The SGM algorithm is faster than global matching algorithms (Hirschmüller and Scharstein, 2009). The SGM algorithm is popular due to the higher efficiency compared with other methods (Alobeid, 2011; Zhu et al., 2014).

The SGM algorithm comprised of four main steps including matching cost computation, cost aggregation, disparity computation and disparity map refinement (Hirschmüller, 2008). The capability of computed costs to match conjugate points has an ambiguity due to equal cost value of neighbor pixels and low entropy in objects. so, the aggregation of matching costs from 8 or 16 paths has been proposed (Hirschmüller, 2008). In this step, matching costs are aggregated based on matching cost and disparity. Equation 1 is used to aggregate the matching costs in each path. Then using equation 2 the costs are aggregated in all paths.

$$L_r(p, d) = C(p, d) + \min \begin{cases} L_r(p-r, d) \\ L_r(p-r, d \pm 1) + p_1 - \min_k L_r(p-r, k) \\ \min_i L_r(p-r, i) + p_2 \end{cases} \quad (1)$$

Where;

p : image location of interest pixel

d : disparity value

$L_r(p, d)$: cost path toward the actual pixel of path

$C(p, d)$: pixel-wise matching cost

$P1$: a small value penalizing disparity changes between neighbouring pixels of one pixel

$P2$: a large value penalizing disparity changes between neighbouring pixels of one pixel

r : actual path

k : pixels in each path

$$S(p, d) = \sum_{r=1}^{8 \text{ or } 16} L_r(p, d) \quad (2)$$

Now, in equation 3 the disparity for each pixel is calculated by minimizing the aggregated cost values. Also, to estimate the disparity in sub-pixel level, a quadratic curve is fitted to the neighboring costs disparity, and the position of the minimum is obtained.

$$D = \min_d S(p, d) \quad (3)$$

After cost aggregation and generating disparity map, mismatching occurs in shadowing areas. The reason is the effect of pixels with lower entropy (such as shadowing areas) to those with higher entropy (such as building roof) to have the same disparity in a local neighbourhood. As shown in Figure 1, this is the reason to extend the roof of the buildings in the disparity map.

The SGM algorithm originally uses the Mean Shift algorithm to refine disparity map in large textureless areas (Hirschmüller, 2008). But this solution fails to solve the ambiguity of shadow effects. Also, some methods were proposed based on segmentation and statistical filtering (such as median of background) to enhance the disparity map in void areas (Bafghi et al., 2016; Krauß and Reinartz, 2010).

In this paper, after shadow detection and computing disparity map from high resolution satellite stereo images, a new method based on morphological filtering and RANSAC plane fitting is proposed to enhance the disparity map in shadow areas. In the following, the proposed method is described and then the result of experiments is presented. The paper will end with the conclusions.

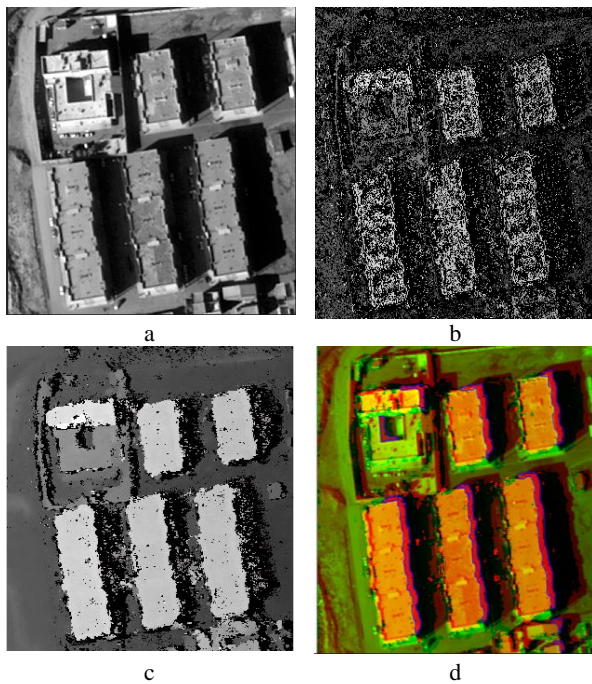


Figure 1. The shadow effect on cost aggregation step (SGM algorithm). a) Panchromatic image; b) computed disparity map without cost aggregation; c) computed disparity map after cost aggregation; d) overlaid disparity map and panchromatic image (the extended roof).

2. PROPOSED METHOD

The proposed method has three main steps, pre-processing, shadow detection, and disparity map refinement. Figure 2 shows the workflow. The details of each part will be discussed in the following.

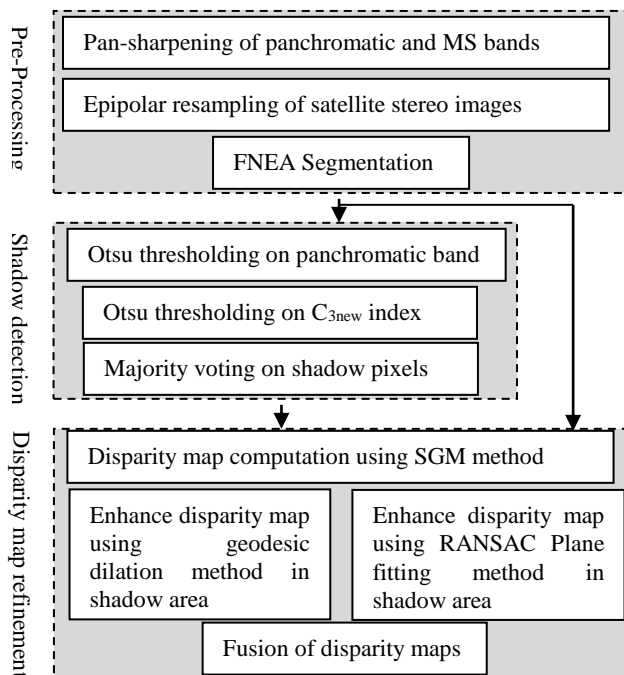


Figure 2. Flowchart of proposed method to enhance the disparity map in SGM algorithm

2.1 Pre-processing

In the pre-processing step, spectral bands are pan-sharpened to enhance the spatial resolution. In the proposed method, the HIS transformation algorithm is used due to its acceptable spatial quality (Strait et al., 2008; Tu et al., 2001).

Epipolar images are rectified images so that each row of the left image corresponds to the same row at the right image (i.e. parallax y is zero). The epipolar images are the input of SGM algorithm (Hirschmüller, 2008). Unlike the images with perspective geometry (Frame Camera), the epipolar geometry of linear array images could not be considered as a straight line. Recently, for linear consideration of epipolar geometry, a new strategy based on image tiling was proposed to produce epipolar images from high resolution satellite stereo images (Tatar et al., 2015b). The panchromatic and pan-sharpened multispectral images are rectified along the epipolar lines using this method.

Then the rectified panchromatic and spectral bands are introduced to Fractal Net Evolution Approach (FNEA) segmentation algorithm (Baatz and Schäpe, 2000; Benz et al., 2004), implemented in eCognition software. This paradigm known as object-based image analysis is used to detect the correct border of Shadow areas in urban area.

2.2 Shadow detection

Detecting shadow areas from high resolution satellite images is a critical role in disparity map refinement. Up to now, a lot of shadow detection methods were proposed, see complete review in (Shahtahmassebi et al., 2013). In this paper, an object-based shadow detection method is used based on our previous work (Tatar et al., 2015a). The method is based on thresholding on the panchromatic band and C_{3new} index to find a binary shadow mask. Then a majority voting (Lam and Suen, 1997) analysis is used to detect shadow areas in object level.

The threshold value is calculated based on Otsu thresholding algorithm (Otsu, 1979). The shadow areas are considered as background and foreground in panchromatic and C_{3new} index respectively. Also, the C_{3new} index was defined by equation 4 (Tatar et al., 2015a).

$$C_{3new} = \tan^{-1} \left(\frac{B}{PAN} \right) \quad (4)$$

Where;

B: Blue band

PAN: Panchromatic band

After detecting shadow objects, connected component analysis (Gonzalez and Woods, 2002) is used to identify shadow regions, and unify them to use as the input for further processing.

2.3 Disparity map refinement

The aggregation step couldn't find the correct matching in shadow and large textureless areas. Therefore, some post processing on disparity map is performed to solve the problem. In this paper, only the effect of shadowing on disparity map will be improved. Two methods of RANSAC (Fischler and Bolles, 1981) plane fitting and geodesic dilation filtering (Arefi and Hahn, 2005) are employed and then the results are integrated to increase the confidence level.

The step by step RANSAC plane fitting is executed as follows.

1. Remove disparities of shadow areas
2. The dilation operator with a 3×3 structure element is applied to the disparity map
3. Using RANSAC algorithm, a plane is fitted to disparities of each shadow region.
4. Using an interpolation process, the shadow pixels getting interpolated disparities

At the same time with RANSAC plane fitting, the geodesic dilation filtering is employed as described below:

1. Detecting non-ground pixels from disparity map using geodesic dilation algorithm
2. Fitting a surface to disparities of ground pixels (estimate background of disparity map)
3. Replace disparities of non-ground pixels by original disparity map

The refined disparity map, the result of RANSAC plane fitting, as well as disparity of geodesic dilation algorithm, permits the determination of false disparities by performing a consistency check. Each disparity of RANSAC plane fitting is compared to its corresponding disparity in the result of geodesic dilation algorithm. The disparity is set to invalid if the difference more than one pixel. Otherwise the disparity is modified by averaging of two disparities.

3. EXPERIMENT AND ANALYSIS

3.1 Dataset

High resolution satellite stereo images from GeoEye-1 are used in the experiments. GeoEye-1 images with panchromatic and multispectral bands are prepared over urban area in Qom city in Iran. Table 1 gives detail information about images, including the acquisition date, Location, and spatial resolutions.

Looking angle	time	Acquired date	Location Lat/Lon	GSD (m)	
				Pan	MS
Afterward	7:20	01/16/2014	N 34.62 °	0.5	2
Forward	7:21		E 50.92°	0.5	2

Table 1: Detail of dataset

3.2 Experimental results

In the pre-processing step, IHS pan-sharpening algorithm used to enhance the spatial resolution of spectral bands which helps to detect shadow areas and create a color point cloud with higher details. After the pan-sharpening, the panchromatic and pan-sharpened spectral bands are resampled along epipolar line. Due to the small size of corresponding stereo image tiles, the epipolar images can be generated by fundamental matrix (Loop and Zhang, 1999; Tatar et al., 2015b). The stereo anaglyph of epipolar images is shown in Figure 3.

In the next step, the epipolar panchromatic and pan-sharpened images are served as input to FNEA segmentation algorithm to produce the image objects. FNEA segmentation algorithm requires tuning the scale parameter, shape and compactness weight coefficients. The value of these parameters has a direct effect on the final segmentation and shadow detection result. The scale parameter, shape and compactness weight coefficients are selected as 80, 0.1 and 0.9 respectively and considered

constant in the experiments based on our previous experiences on the same dataset (Tatar et al., 2015b; Tatar et al., 2016).

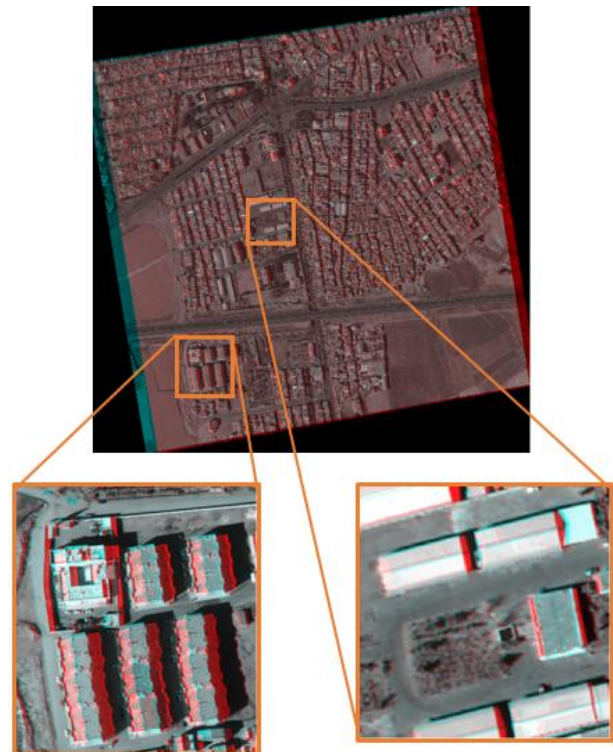


Figure 3. Stereo anaglyph of the generated epipolar images

Recognition of shadow objects in the proposed method depends on the detection of suspected shadow pixels. Thresholding on the panchromatic band and the C_{3new} index are employed to this end. Thresholding on the panchromatic band detects a mixture of dark pixels including shadow pixels, asphalt road, and dark roofs. In the next step, thresholding on C_{3new} index improves the accuracy of suspect shadow pixels by omitting the false shadow pixels among the suspicious shadow pixels.

Then the results of pixel-level shadow detection and FNEA image segments are integrated through the majority voting to recognize shadow areas in object level. This will improve the accuracy with omitting salt and pepper noise in suspect shadow detection. The threshold percentage of the shadow pixels in each image segment is needed to make decisions here. From previous studies, the best majority threshold is selected as 20% (Tatar et al., 2015a; Tatar et al., 2016). Figure 4 shows the comparison of pixel and object level shadow detection results.

After generating epipolar images for corresponding stereo image tiles, the SGM algorithm is applied to generate disparity map. As mentioned earlier, matching cost function, and penalty of P_1 and P_2 are critical inputs for SGM algorithm. The Census transformation with 7×9 kernel size is selected as the cost function. Also, the value of P_1 and P_2 are obtained 12 and 30 respectively in the experiment. The noises and outliers are removed using a median filter.

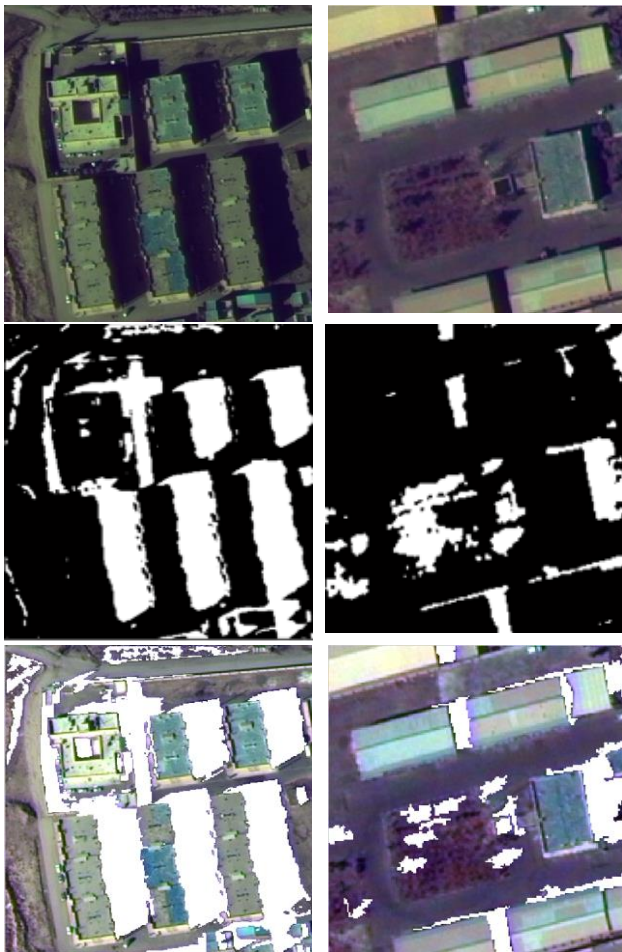


Figure 4. Shadow detection results. Top: pan-sharpened original images; middle: suspected shadow areas; bottom: overlaid object-based shadow areas on original images

Cost aggregation in SGM algorithm is sensitive to entropy values and it caused a shift in disparity value of shadow areas to get closer to those of neighbouring buildings. To solve this problem, as described in section 2.3, the disparity map refinement is applied to the disparity map. The generated disparity map of SGM algorithm and enhanced disparity map are shown in Figure 5.

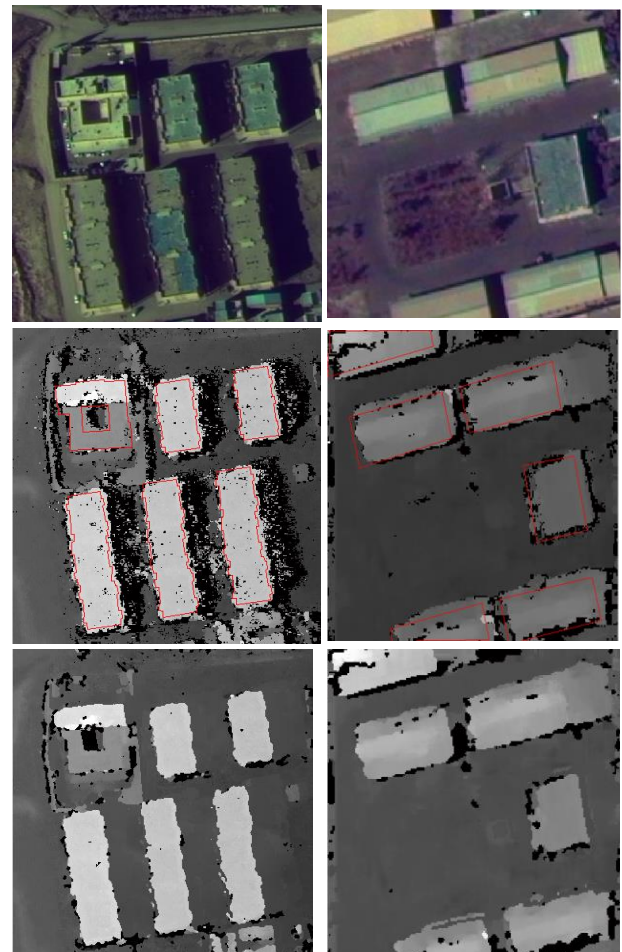


Figure 5. Disparity map refinement. Top: original epipolarized left image; middle: Computed disparity map by SGM algorithm (the border of man-made objects drawn in red line); bottom: enhanced disparity map (proposed method)

Also to visual inspection of shadow effect on SGM algorithm, the corresponding points and images were intersected and color point cloud is generated (Figure 6). Due to the existence of rational polynomial coefficients (RPCs) for satellite stereo images, the corresponding points are intersected by RPCs in the object coordinate system.

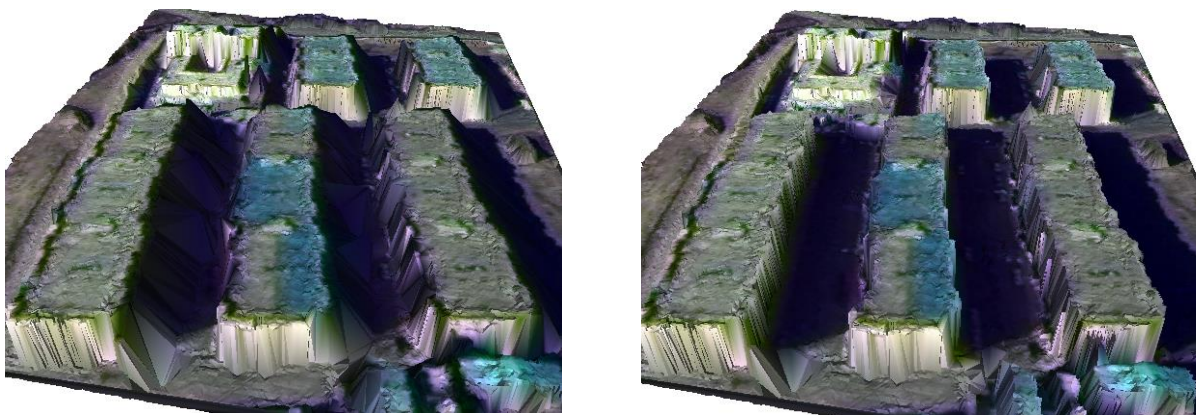


Figure 6. 3D reconstruction of high resolution satellite stereo images. Left: 3D reconstruction with shadow effect (standard SGM algorithm); right: 3D reconstruction after disparity map refinement (proposed method).

3.3 Evaluation and Discussion

In this section, the result of epipolar resampling, shadow detection, and disparity map refinement are evaluated. Y-parallax values for corresponding points in epipolar images considered as evaluation measure to the results of epipolar resampling. The Y-parallax value of measured corresponding points is shown in Figure 7. The result of experiments in Figure 3 and 7 demonstrates that the epipolar images are generated with sub-pixel accuracy.

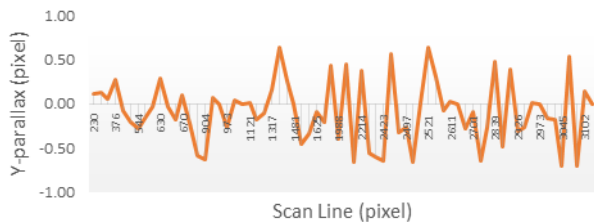


Figure 7. Y-parallax of corresponding points on the generated epipolar images

The confusion matrix is computed to evaluate the result of shadow detection and disparity map refinement. The confusion matrix for generation of two classes is shown in Table 2.

		Detected /generated	
		Class 1	Class 2
Reference	Class 1	TP	FP
	Class 2	FN	TN

Table 2: Confusion matrix for evaluation of results

The confusion matrix in shadow detection includes: true positive (TP) number of correctly classified shadow pixels, false positive (FP) number of wrongly classified non-shadow pixels as shadow, false negative (FN) number of shadow pixels which detected as non-shadow, and true negative (TN) number of non-shadow pixels classified correctly.

Also, the confusion matrix is created for evaluation of disparity map refinement and computed disparity map by SGM algorithm. In this evaluation, the confusion matrix includes:
 TP: number of correctly matched pixels
 FP: number of mismatched pixels
 FN: number of conjugate pixels which considered as occlusion
 TN: number of occlusions truly found.

Completeness, correctness and F-measure criterions obtained from confusion matrix as below are used to evaluate the results.

$$Completeness = 100 \times \frac{TP}{TP+FN} \quad (5)$$

$$Correctness = 100 \times \frac{TP}{TP+FP} \quad (6)$$

$$F - measure = 100 \times \frac{2 \times TP}{2 \times TP + FN + FP} \quad (7)$$

To evaluate the shadow detection results, 98 image objects in shadow class and 118 image objects in non-shadow class are selected manually. Compactness, correctness, and F-measure were calculated based on generated confusion matrix. Table 3 contains the result of object-based shadow detection. In this table, the F-measure is used as quantitative criteria due to the big gap between correctness and completeness.

Completeness	Correctness	F-measure
100%	89%	94%

Table 3: Statistical analysis of object-based shadow detection

To evaluate the performance of proposed disparity map refinement method, 260 corresponding points are measured manually. The disparity value of measured points was compared to the disparity computed by SGM algorithm and the proposed method. Evaluation results for both results are presented in Table 4.

Method	Completeness	Correctness	F-measure
SGM algorithm	48%	71%	57%
Proposed method	89%	94%	92%

Table 4: Statistical analysis of computed disparity map by SGM algorithm and enhanced by proposed method

To evaluate the computed disparity map, completeness, correctness, and F-measure are used. It should be noted that the higher value of correctness, shows the non-occlusion pixels matched correctly. In the other word, the building roofs and land cover are matched correctly. According to the results presented in Table 4, lower correctness value in the case of SGM algorithm indicates the number of shadow pixels (wrongly matched) which added to all matched pixels. Due to the enhanced disparity map in shadow areas, the wrongly matched pixels are removed from all matched pixels. This will cause an increase in the correctness value in the case of proposed method.

Also, if completeness has a higher value, indicates the occlusion and shadow pixels are not matched as non-occlusion and roof pixels respectively. Quantitative evaluation in Table 4, shows the lower completeness in the case of original SGM algorithm. The reason is that the shadow pixels matched as roof pixels. However, in the case of proposed method, the completeness is increased due to disparity refinement in shadow areas.

4. CONCLUSION

In this paper, we investigate the effect of shadow area on SGM algorithm in the case of high resolution satellite stereo processes. The SGM algorithm causes shadow pixels have a disparity equal to the disparity of buildings roof. To solve this problem, FNEA segmentation was employed to detect shadow areas accurately in object level. Then a new method based on RANSAC plane fitting and morphological filtering used to enhance the disparity map in shadow areas. The result of experiments demonstrates that the proposed object-based disparity map refinement significantly increased the result of original SGM algorithm in shadow areas.

The proposed method in this paper focuses on disparity map refinement in shadow areas. For other areas such as mirrored surfaces and non-lamebrain surfaces, an adaptive disparity map refinement method seems to work. Also, the numerical assessment of the results was done by some corresponding points which measured manually. Creating a dataset with ground truth and evaluate the proposed method with ground truth will follow by the authors.

REFERENCES

- Alobeid, A., 2011. Assessment of matching algorithms for urban DSM generation from very high resolution satellite stereo images. *Fachrichtung Geodäsie und Geoinformatik der Leibniz-Univ., Hannover, Germany*.
- Arefi, H., Hahn, M., 2005. A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.
- Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation, in: Strobl, J. (Ed.), *Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000*, Karlsruhe, Herbert Wichmann Verlag, pp. 12-23.
- Bafghi, Z.G., Tian, J., d'Angelo, P., Reinartz, P., 2016. A New Algorithm For Void Filling in A DSM From Stereo Satellite Images in Urban Areas. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 3.
- Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of photogrammetry and remote sensing* 58, 239-258.
- Brown, M.Z., Burschka, D., Hager, G.D., 2003. Advances in computational stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 25, 993-1008.
- Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24, 381-395.
- Gonzalez, R.C., Woods, R.E., 2002. *Digital image processing*.
- Hirschmüller, H., 2008. Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 30, 328-341.
- Hirschmüller, H., Scharstein, D., 2009. Evaluation of stereo matching costs on images with radiometric differences. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31, 1582-1599.
- Krauß, T., Reinartz, P., 2010. Enhancement of dense urban digital surface models from VHR optical satellite stereo data by pre-segmentation and object detection, *Int Arch Photogramm Remote Sens Spat Inf Sci XXXVIII (Part 1)*, on CDROM.
- Lam, L., Suen, S., 1997. Application of majority voting to pattern recognition: an analysis of its behavior and performance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 27, 553-568.
- Loop, C., Zhang, Z., 1999. Computing rectifying homographies for stereo vision, *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on. IEEE, Fort Collins, CO, USA*.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on System, Man, and Cybernetics SMC-9*, 5.
- Scharstein, D., Szeliski, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision* 47, 7-42.
- Shahtahmassebi, A., Yang, N., Wang, K., Moore, N., Shen, Z., 2013. Review of shadow detection and de-shadowing methods in remote sensing. *Chinese Geographical Science* 23, 403-420.
- Strait, M., Rahmani, S., Merkurev, D., 2008. Evaluation of pan-sharpening methods. *UCLA Department of Mathematics, Research Experiences for Undergraduates (REU2008)*.
- Tatar, N., Saadatseresht, M., Arefi, H., Hadavand, A., 2015a. A New Object-Based Framework to Detect Shadows in High-Resolution Satellite Imagery Over Urban Areas. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40, 713.
- Tatar, N., Saadatseresht, M., Arefi, H., Hadavand, A., 2015b. Quasi-Epipolar Resampling of High Resolution Satellite Stereo Imagery for Semi Global Matching. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40, 707.
- Tatar, N., Saadatseresht, M., Arefi, H., Hadavand, A., 2016. Developing New Index to Object-Based Shadow Detection from High Resolution Satellite Images over Urban Area. *Journal of Geomatics Science and Technology* 5, 11-21.
- Tu, T.-M., Su, S.-C., Shyu, H.-C., Huang, P.S., 2001. A new look at IHS-like image fusion methods. *Information Fusion* 2, 10.
- Zhu, K., der von der Ingenieur fakultät, V.A., Geo, B., Vermessungswesen, U.B.-u., 2014. *Dense Stereo Matching with Robust Cost Functions and Confidence-based Surface Prior*. *Universitätsbibliothek der TU München*.