Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography
Vol. 37, No. 5, 315-322, 2019
https://doi.org/10.7848/ksgpc.2019.37.5.315

ISSN 1598-4850(Print)
ISSN 2288-260X(Online)
Original article

Topographic Information Extraction from Kompsat Satellite Stereo Data Using SGM

Jang, Yeong Jae [1] · Lee, Jae Wang [2] · Oh, Jae Hong [3]

Abstract

DSM (Digital Surface Model) is a digital representation of ground surface topography or terrain that is widely used for hydrology, slope analysis, and urban planning. Aerial photogrammetry and LiDAR (Light Detection And Ranging) are main technology for urban DSM generation but high-resolution satellite imagery is the only ingredient for remote inaccessible areas. Traditional automated DSM generation method is based on correlation-based methods but recent study shows that a modern pixelwise image matching method, SGM (Semi-Global Matching) can be an alternative. Therefore this study investigated the application of SGM for Kompsat satellite data of KARI (Korea Aerospace Research Institute). Firstly, the sensor modeling was carried out for precise ground-to-image computation, followed by the epipolar image resampling for efficient stereo processing. Secondly, SGM was applied using different parameterizations. The generated DSM was evaluated with a reference DSM generated by the first pulse returns of the LIDAR reference dataset.

Keywords : Stereo, Satellite, Kompsat, DSM, SGM, Matching

1. Introduction

DSM (Digital Surface Model) is a digital 3D representation of topographic features and terrains in geographic coordinates. DSM is widely used for map generation, flight simulation, urban planning and monitoring. Conventionally DSM was created from aerial photos and satellite images. Terrain contours and building boundaries are extracted from those stereo images and the elevation information is interpolated for DSM generation. But this approach requires human operators and labor-intensive, not to mention the cost. Therefore an automated approach such as dense stereo matching has been required and studied (Alobeid *et al*., 2010).

Satellite data for city-level DSM generation had been limited because of the relative low resolution but the high resolution satellites such a s IKONOS, QuickBird, WorldView enabled the object-level feature extraction. Korea also launched Kompsat-2, 3 that show sub-meter level spatial resolution with up to 19km coverage in single scene. Kompsat-3 has a capability of stereo image acquisition in single trajectory such that studies for urban level DSM generation based on Kompsat-3 are required.

DSM extraction from stereo images typically consists of precise sensor modeling, epipolar image resampling, automated stereo matching, and ground coordinate reconstruction (Oh and Lee, 2016). Aforementioned processes except for the stereo matching show high accuracy based on rigorous mathematical modeling. But the stereo matching produces a lot of mismatches due to the geometric and radiometric differences between the stereos, and occlusions. In this case, the quality of topographic information is highly reduced. The popular stereo matching

Received 2019. 09. 18, Revised 2019. 09. 30, Accepted 2019. 10. 10
1) Member, Dept. of Civil Engineering, Korea Maritime and Ocean University (E-mail: 93l73444@kmou.ac.kr)
2) Dept. of Civil Engineering, Korea Maritime and Ocean University (E-mail: jnso5072@outlook.kr)
3) Corresponding Author, Member, Dept. of Civil Engineering, Korea Maritime and Ocean University (E-mail: jhoh@kmou.ac.kr)

techniques for DSM generation are NCC (Normalized Cross Correlation matching), LSM (Least Squares Matching), and SGM (Semi-Global Matching). NCC uses the digital number of pixels in reference and target images to compute the similarity and to locate the conjugate point. To overcome the brightness difference between the reference and target images, image normalization process is necessary. LSM iteratively adjusts the conjugate point locations to minimize the brightness difference between the reference and the target. This technique is known to show more accurate image coordinates compared to NCC (Poon *et al.*, 2007; Aguilar *et al.*, 2014), but if the initial image coordinates for the iteration is not accurate, the solution does not converge. SGM is first proposed by Hirschmuller (2008) using MI (Mutual Information) between the stereo set for the cost. In SGM, the matching cost is aggregated along multiple paths with assigning penalties to sudden disparity changes, and the pixel with minimum cost is selected for the conjugate point. SGM is known to be invariant to noise, scale, rotation, and brightness and has high potential for precise topographic information generation from stereo satellite images. Therefore, some researches applied SGM for DSM generation from satellite stereo data (Ghuffar, 2016; Gong and Fritsch, 2018).

This study applied SGM to Kompsat-3 stereo data and carried out accuracy and precision assessment of ground target reconstruction. We applied different parameterizations such as search region sizes and penalties in SGM and their effects are analyzed using aerial LiDAR (Light Detection And Ranging) data that show typical elevation accuracy of 15 cm for the reference.

This paper is structured as below; Chapter 2 introduces preprocessings such as the satellite sensor modeling and the epipolar image resampling. In addition, SGM is also explained. Chapter 3 presents the experimental results and analysis for Kompsat-3 data, followed by the conclusion in Chapter 4.

2. Methodology

2.1 Satellite Sensor Modeling

Recent high-resolution earth observing satellites have a capability of acquiring overlapping images in single trajectory by changing the pitch angle as shown in Fig. 1. The images acquired in this mode enable higher quality topographic information by minimizing the radiometric difference and increasing similarity between the stereo data. However the satellite data must be accurately preprocessed to ensure the geometric consistency between the data as well as increasing geographic coordinates of the final product. This step is called 'sensor modeling'. The sensor modelling of high-resolution satellite is carried out using which forms the nonlinear equation as Eq. (1), an image coordinates (s,l) is computed from a given ground coordinates (U,V,W). For a precise stereo processing, the refinement parameters A, B in the equation must be estimated to remove the geometric inconsistency between a stereo data set.

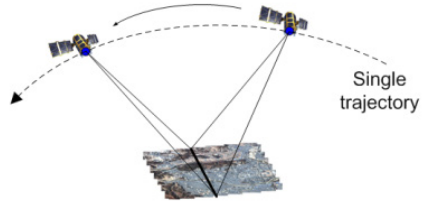


Fig. 1. Stereo satellite data acquisition in single trajectory

$$\begin{aligned} l + A_0 + A_1 l + A_2 s &= F_1(U,V,W)/F_2(U,V,W) \\ s + B_0 + B_1 l + B_2 s &= F_3(U,V,W)/F_4(U,V,W) \end{aligned} \quad (1)$$

where, l, s are line and sample coordinates and F_i are third-order polynomial functions of object space coordinates U, V and W. $\mathrm{A}_0, \mathrm{A}_1, ..., \mathrm{B}_2$ describe an refinement parameters.

2.2 Epipolar Image Resampling

Efficient and fast stereo processing and 3D display require the epipolar geometry analysis and epipolar image resampling. But the stereo geometry of a pushbroom satellite sensor is quite different from that of a frame camera (Oh *et al.*, 2010). Fig. 2 shows the resampling the epipolar curve points to satisfy the epipolar image conditions, which is zero y-disparity (no difference in image row coordinates) and the linear relationship between the x-disparity and the ground elevation. Note that zero y-disparity is implemented by assigning a constant row coordinate value to each epipolar curve pair in both images. The linear relationship between the x-disparity (difference in

image column coordinates) and the ground elevation can be achieved by setting the constant interval in-between the epipolar curve points..

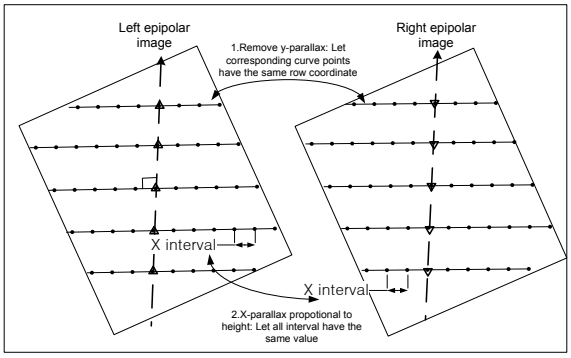


Fig. 2. Epipolar resampling method from the generated epipolar curve image points (Oh *et al.*, 2010)

2.3 Stereo Matching

The flow of SGM is shown in Fig. 1. I_b and I_m are the base and match images, respectively, and D_{init} is initial disparity information. First I_m is warped using the initial disparity information for $f_D(I_m)$, then Mutual Information is computed (II-A). Next step is the cost calculation and cost aggregation based on the Mutual Information for each pixel and disparity (II-B).

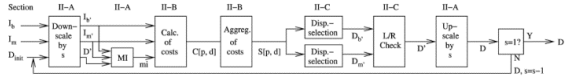


Fig. 3. Flowchart of SGM matching (Hirschmuller, 2008)

Mutual Information is computed from entropy and joint entropy as Eqs. (2)-(4). When the base image and matched images are well registered, the joint entropy is minimized (Maes, 1997) because one image can be predicted by the other image. This leads to higher mutual information.

$$mi_{I_1,I_2}(i,\mathrm{k}) = h_{I_1}(i) + h_{I_2}(k) - h_{I_1,I_2}(i,k) \tag{2}$$

$$h_I(i) = -\frac{1}{n}\log\left(P_I(i)\otimes g(i)\right)\otimes g(i) \tag{3}$$

$$h_{I_1,I_2}(i,k) = -\frac{1}{n}\log\left(P_{I_1,I_2}(i,k)\otimes g(i,k)\right)\otimes g(i,k) \tag{4}$$

where, h_I is an entropy, h_{I_1,I_2} is the joint entropy, P is the probability distribution, g is a Gaussian smoothing, mi is the mutual information.

Matching cost is negative to the Mutual Information as shown in Eq. (5). In other words, higher Mutual Information (lower joint entropy of well registered images) leads to lower matching cost. This cost computation is carried out for all pixels in the disparity range such that a 3D cube of matching cost matrix is generated.

$$C(p,d) = -mi_{I_b,f_D(I_m)}\left(I_{bp}, I_{mq}\right) \tag{5}$$

The cost aggregation step refines the matching cost result considering pixel location change r and disparity d. Aggregation equation for a direction r is shown in Eq. (6) depicting accumulating the matching cost along the direction. All direction computation is computationally intensive such that cost aggregations along 8 or 16 directions are carried out.. Fig. 4 shows the case of 16 directions where a line of minimum cost in 3D disparity cube in the left.

$$L_r(p,d) = C(p,d) + min\begin{cases} L_r(p-r,d) \\ L_r(p-r,d+1)+P_1 \\ L_r(p-r,d-1)+P_1 \\ min_i L_r(p-r,i)+P_2 \end{cases} - min_k L_r(p-r,k) \tag{6}$$

where, r is a direction, P_1 is penalty for small changes in neighborhood, P_2 is adaptive penalty for large changes in neighborhood, L_r is the aggregated cost.

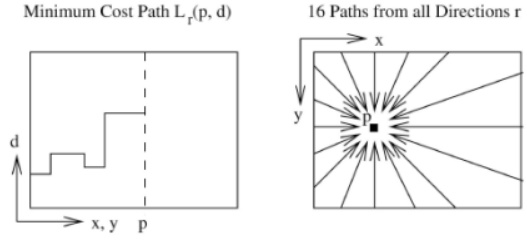


Fig. 4. Matching cost aggregation (Hirschmuller, 2008)

3. Experiment

3.1 Test Kompsat-3 Stereo Satellite Data

Test data are Kompsat-3 stereo data acquired in 18 Mar 2014 over Daegu, Korea as shown in Fig. 5 with ground control points locations in triangle symbols. GSD (Ground

Sampling Distance) is about 0.8 meters and single image size is 24000x24000 pixels showing ground coverage of 19kmx19km. The ground control points were used for accurate sensor modeling and affine-based bias-compensation (Fraser and Hanley, 2005). The geographic accuracy of the original data was 34~40 pixels but the bias-compensation improved the accuracy up to about one pixel in RMSE (Root Mean Square Error).

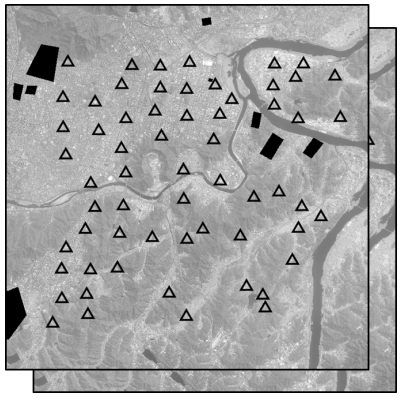


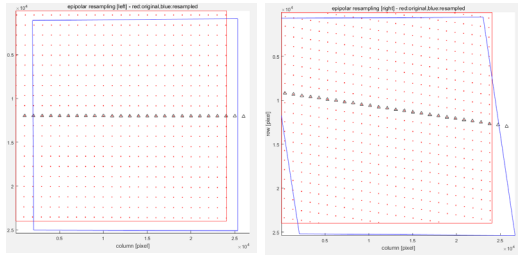
Fig. 5. Kompsat-3 stereo images with control points distribution

3.2 Epipolar Image Resampling

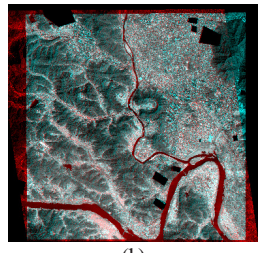
Epipolar image resampling was carried out using the bias-compensated sensor modeling information. First epipolar points were extracted over the entire image region for both images as shown in Fig. 6(a) with red dots. The red box is the boundary of original images, the blue boundary shows the epipolar resampled image boundary. Fig. 6(b) shows the epipolar resampled images in anaglyph.

3.3 Stereo Matching

We selected a test region for SGM matching from the epipolar resampled images as presented in Fig. 7(a). The test region includes factories and flat terrains with elevation range of 50m~70m. Before the image matching, we removed noises by applying the median Filter. Fig. 7(b) shows the aerial LiDAR DSM that is used for elevation accuracy assessment. The data well shows boundaries of buildings but we can identify a small artifact in the dotted circle that has much lower elevation values even lower than flat neighbor terrains. The horizontal direction of the epipolar images in Fig. 7(a) is the satellite's trajectory such that the orientation of the images are not same with LiDAR DSM in Fig. 7(b). Note that the arrow indicates same building object.



(a)



(b)

Fig. 6. Epipolar resampling (a) before and after the resampling - red: original image boundary, blue: resampled boundary (b) resampled images - anaglyph

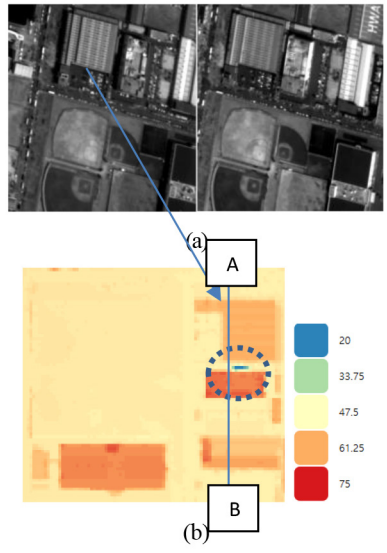


Fig. 7. Stereo matching test area (a) stereo satellite data, (b) aerial LiDAR data with elevation

3.3.1 Disparity Parameterization

Disparity is a column coordinates difference between epipolar resampled images. SGM requires a disparity range for the input. We tested five different disparity ranges such as [1 ~ 96], [16 ~ 96], [32 ~ 96], [48 ~ 96], [64 ~ 96] to check their effects. Note that the actual disparity range of the test region is about 30 ~ 80 pixels. We fixed penalty parameters as P1=2,904, P2=11,616 for all the cases. Fig. 8 shows the matching results in disparity map for the five cases. The disparity range [32~96] that is closest to actual range [30~80] shows the least noisy result as shown in Fig. 8(c) but the other cases describe the building boundaries well except for Fig. 8(e). Test case [64~96] is outside of the actual disparity range failed to extract the topographic information as shown in Fig. 8(e). Therefore, even disparity ranges are set larger than the actual range, SGM could produce acceptable matching results as long as the disparity range covers the actual disparity range. The disparity results were used for the ground reconstruction for DSM generation. Then the elevation of each DSM was compared to the aerial LiDAR data. Fig. 9 shows the elevation accuracy of each DSM. The disparity range [32~96] showed the lowest error (highest accuracy) but the range [64 ~ 96] showed the elevation error more than 25 meters.

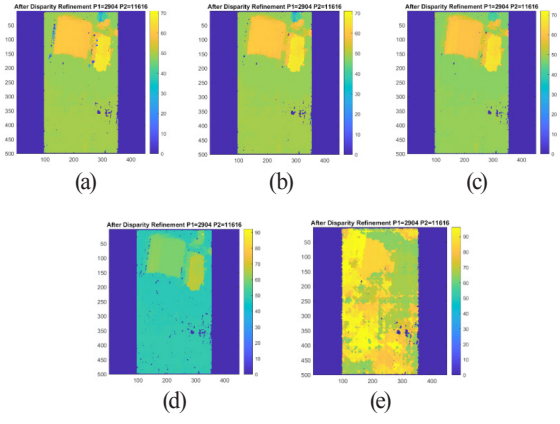


Fig. 8. Disparity map for each SGM test with different disparity ranges for input: (a) [1 ~ 96], (b) [16~ 96], (c) [32 ~ 96], (d) [48 ~ 96], (e) [64 ~ 96]

We compared the topographic profiles of satellite DSM and LiDAR DSM to check precision of the object shape extraction. We generated the profiles along A-B section as depicted in Fig. 7(b). In Fig. 10, LiDAR DSM and satellite DSM are plotted in regular lines and dotted lines, respectively. Note that significantly lower elevation is observed between the building objects in LiDAR DSM as we already pointed out in Fig. 7. Except for the part, most test cases did not fail to describe the building height except for Fig. 10(e). Also two DSMs do not show significantly large differences except for the part as most cases except for the case [64~96] show errors ranges 3~4 meters in Table 1. Considering the spatial resolution of Kompsat-3, the elevation accuracy is about 4~5 pixels level. But most large elevation errors occur along the building boundaries due to the geographic coordinate mismatches, and the elevation error over a flat region is much lower.

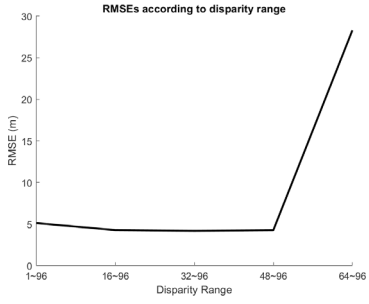


Fig. 9. Elevation accuracy of each disparity range case in RMSE

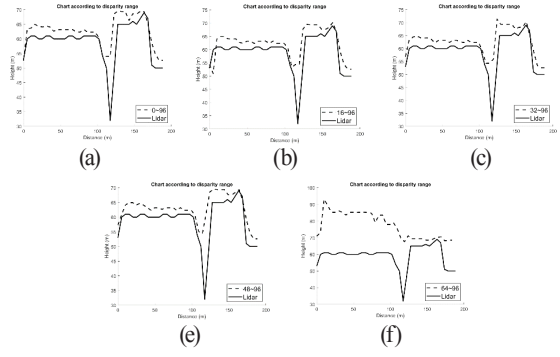


Fig. 10. Topographic profile comparison between satellite DSM and aerial LiDAR DSM : (a) 0 ~ 96, (b) 16~ 96, (c) 32 ~ 96, (d) 48 ~ 96, (e) 64 ~ 96

Table 1. Elevation errors along the profile for different disparity ranges

Disparity range	Elevation error
0~96	3.76
16~96	4.00
32~96	3.33
48~96	3.43
64~96	21.95

3.3.2 Penalty Parameterization

We tested different penalty parameters for SGM matching. Penalty is a parameter that controls disparity smoothness and there are two penalties, P1 and P2. P1 is a penalty used to control the disparity change of neighboring pixels such as +1, -1 pixels. P2 is used for larger disparity changes than +1, -1 pixels. In other words, penalty parameters are used to suppress the abrupt disparity changes. We tested five different penalty sets as Eq. (7) and Eq. (8). We increased SADWindowSize such as 3, 5, 7, 9, 11 to set the penalty parameters (OpenCV team, 2019). And fixed disparity range as [30~80] for all the cases.

$$P1 = 8 * 3 * P * SADWindowSize^2 \quad (7)$$

$$P2 = 32 * 3 * SADWindowSize^2 \quad (8)$$

Fig. 11 shows the matching results in disparity map for the five cases. Lower penalty case such as Fig. 11(a) could not describe the building object well but the noises reduce as penalty increases. Fig. 11(b) shows much less noise than Fig. 11(a) but more penalty increase does not affect much. Then the elevation of each DSM was compared to the aerial LiDAR data and the elevation accuracy of each DSM was plotted in Fig. 11. The graph shows that the error decreases as the penalty increases. The elevation error significantly reduced from case [216/864] to case [600/2400] but meaningful error decrease was not observable in other cases because the elevation change 0.1m is not significant considering the satellite image's resolution.

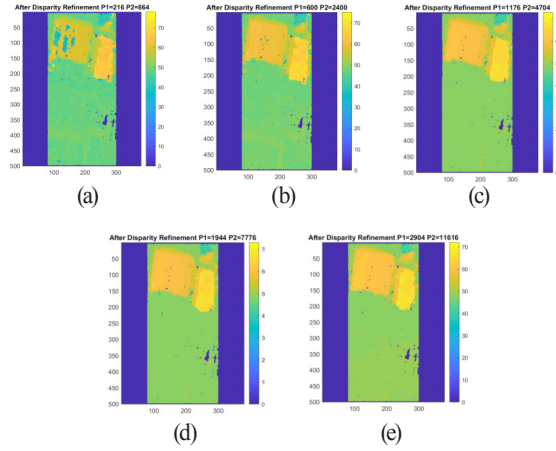


Fig. 11. Disparity map for each SGM test with different penalty parameters for input: (a) 216/864 (Penalty1/ Penalty2), (b) 600/2400, (c) 1176/4704, (d) 1944/7776 (e) 2904/11616

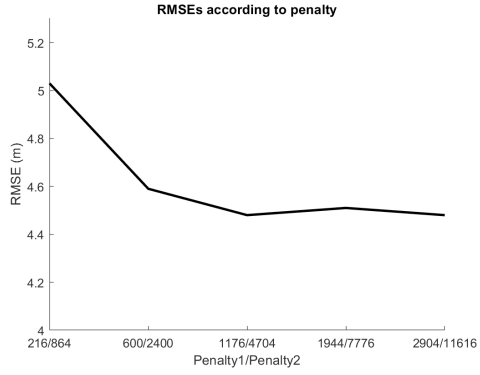


Fig. 12. Elevation accuracy of each penalty case in RMSE

Fig. 13 shows the profile comparison. Significant noises are observed in building region in Fig. 13(a) but other cases described the building objects even though the elevation is overestimated than the reference by 3~4 meters.

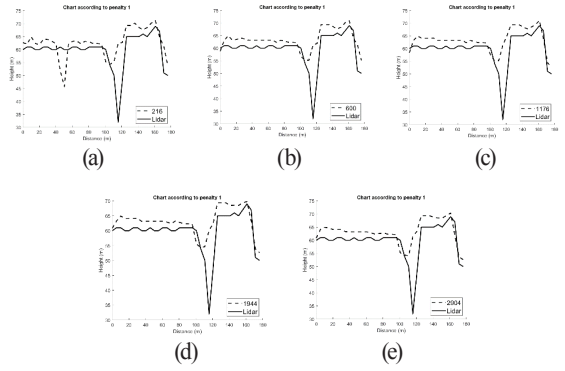


Fig. 13. Topographic profile comparison between satellite DSM and aerial LiDAR DSM: (a) 216/864 (Penalty1/ Penalty2), (b) 600/2400, (c) 1176/4704, (d) 1944/7776 (e) 2904/11616

Table 2. Elevation errors along the profile for different penalty parameters

Penalty1 / Penalty2	Elevation error
216 / 864	4.57
600 / 2400	4.13
1176 / 4704	3.99
1944 / 7776	3.85
2904 / 11616	3.98

Fig. 14 presents LiDAR DSM and Kompsat-3 DSM from SGM matching in 3D view for comparison.

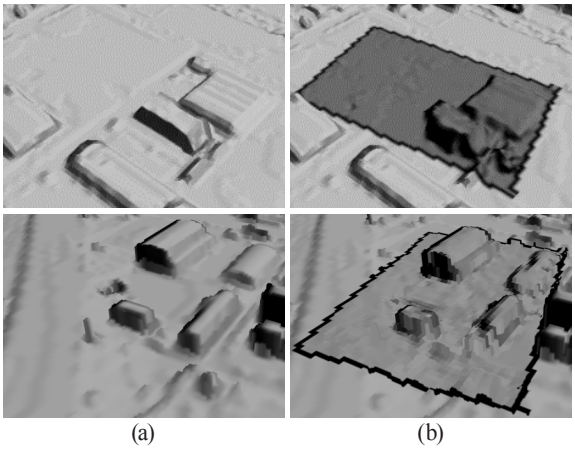


Fig. 14. 3D view of DSM (a) LiDAR DSM, (b) Kompsat-3 DSM from SGM matching

4. Conclusion

The study tested SGM matching for Kompsat-3 stereo satellite data and analyzed the elevation accuracy and precision of topographic information extraction. Kompsat-3 stereo satellite data acquired in single trajectory were preprocessed with the accurate sensor modeling and epipolar image resampling. Then we applied different parameterization of disparity range and penalties as inputs for SGM. The result disparity maps and ground reconstructed DSM are compared to accurate aerial LiDAR DSM for the accuracy assessment. As results of the disparity range test, increasing the range tends to produce more matching noise. But even disparity ranges are set larger than the actual range, SGM could produce acceptable matching results as long as the disparity range covers the actual disparity range. Lower penalty tends to allow abrupt disparity changes resulting incomplete building object description. Increasing penalty parameters can increase the elevation accuracy of the DSM but no meaningful accuracy increase is not observable from certain values. In the experiment, we could obtain 3~4 meters of elevation accuracy from the tested Kompsat-3 stereo data. For practical DSM generation, it is recommended to try different disparity ranges and penalty parameters. For the future study, more SGM parameterization experiments will be carried out for other satellite data such as WorldView series, Pleiades, and Korea's compact advanced satellite that will be launched in 2020. In addition, multiple image matching techniques will be studied to increase precision and accuracy of urban DSM generation.

Acknowledgements

This work was supported by the Korea Maritime And Ocean University Research Fund.

References

Aguilar, M., Saldana, M., and Aguilar, F. (2014), Generation and quality assessment of stereo-extracted DSM from GeoEye-1 and WorldView-2 imagery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52, No. 2, pp.

1259-1271.

Alobeid, A., Jacobsen, K., and Heipke, C. (2010), Comparison of matching algorithms for DSM generation in urban areas from IKONOS imagery, *Photogrammetric Engineering & Remote Sensing*, Vol. 76, No. 9, pp. 1041-1050.

Fraser, C.S. and Hanley, H.B. (2005), Bias-compensated RPCs for sensor orientation of high-resolution satellite imagery, *Photogrammetric Engineering & Remote Sensing*, ASPRS, Vol. 71, No. 8, pp. 909-915.

Ghuffar, S. (2016), Satellite stereo based digital surface model generation using semi global matching in object and image space, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. III-1, 12-19 July, Prague, Czech Republic, pp. 63-68.

Gong, K. and Fritsch, D. (2018), Point cloud and digital surface model generation from high resolution multiple view stereo satellite imagery, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-2, 4-7 June, Riva del Garda, Italy, pp. 363-370.

Hirschmuller, H. (2008), Stereo processing by semiglobal matching and mutual information, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 2, pp. 328-341.

Maes, F. (1997), Multimodality image registration by maximization of mutual information, *IEEE Transactions on Medical Imaging*, Vol. 16, No.2, pp.187-198.

Oh, J.H. and Lee, C.N. (2016). Extraction of digital elevation model using stereo matching with slope-adaptive patch transformation, *KSCE Journal of Civil Engineering*, Vol. 20, No. 7, pp. 2902-2909.

Oh, J.H., Lee, W.H., Toth, C.K., Grejner-Brzezinska, D.A., and Lee, C.N. (2010), A piecewise approach to epipolar resampling of pushbroom satellite images based on RPC, *Photogrammetric Engineering & Remote Sensing*, Vol. 76, No. 12, pp. 1353-1363.

OpenCV team (2019), Class StereoSGBM, *OpenCV Doc*, https://docs.opencv.org/java/2.4.9/org/opencv/calib3d/StereoSGBM.html (last date accessed: 18 September 2019).

Poon, J., Fraser, C.S., and Chunsun, Z. (2007), Digital surface models from high resolution satellite imagery, *Photogrammetric Engineering & Remote Sensing*, Vol. 72, No. 11, pp. 1225-1231.