

3D SURFACE RECONSTRUCTION FROM MULTI-VIEW AND MULTI-DATE GOOGLE EARTH SATELLITE IMAGES WITH 3D HOMOGRAPHY-BASED PROJECTIVE RECONSTRUCTION

M. J. Lee¹, S. Y. Park¹ *

¹ School of Electronics and Electrical Engineering, Kyungpook National University, Daegu 41566, South Korea

KEY WORDS: Multi-view stereo matching, Multi-date satellite images, 3D projective reconstruction, Google Earth.

ABSTRACT:

In this paper, we propose the 3D surface reconstruction scheme using multi-view and multi-date Google Earth (GE) satellite images. Multi-view stereo matching (MVS) scheme is one of the methods for reconstructing dense 3D surface based on multi-view images and corresponding camera pose geometry. If many views are input, MVS can estimate the disparity (depth) by matching pixels. However, the common users are not always possible to obtain both multi-view satellite images and the camera geometry (such as Rational Polynomial Camera) in various earth regions. Instead, the GE provides multi-view and multi-date satellite images of earth regions. Therefore, the goal of the proposed method is to perform a 3D surface reconstruction using the GE satellite image. We suppose that the GE satellite image is a pinhole camera model, and the camera pose geometry is estimated using the perspective projection model (PPM) based structure from the motion (SfM) method. Then the 3D surface is reconstructed and fusion using the MVS method. However, the GE satellite image is a transformed pseudo-orthoimage for integration into the raster image. For this reason, the camera pose geometry is inaccurately computed in the SfM process. Thus, the high-rise structures in the reconstructed 3D surface are distorted (distorted hexahedral 3D space). Importantly, the satellite image is a weak PPM and it can express the orthographic projection model. Therefore, we compute 3D homography for transforming between distorted hexahedral space to orthographic cuboid space. Then, the distorted 3D surface is transformed using a projective reconstruction based on 3D homography. The transformed 3D surface has the correct shape in the orthographic projection model. The advantage of the proposed method is that the 3D surface of various earth regions is reconstructed using simply accessible GE satellite images. And the transformed 3D surface is reconstructed into orthographic projection model space, thus the orthoimage can be generated using projection.

1. INTRODUCTION

The large-ground 3D surface is used in several applications such as remote sensing, simulation, and 3D mapping. The stereo matching scheme is one of the methods for obtaining 3D data (Hirschmuller, 2005). The theory of the stereo matching method is that the disparity (depth) is estimated using pixel matching in two or more input images. Since the disparity in all pixels of the input image is estimated, it can perform dense 3D reconstruction. In addition, it can estimate the accurate disparity as the multi-views (more than two views) (Kendall, et al., 2017, Im et al., 2019, and Xu, et al., 2020). The MVS scheme of the general camera model reconstructs 3D data using both the images of various view directions and corresponding camera pose geometry (camera intrinsic, camera extrinsic parameters). In the case of multi-view satellite images, the camera pose geometry is the position of the satellite, and it is expressed as Rational Polynomial Camera (RPC). However, in the stance of common users, it is not always possible to obtain both multi-view satellite images and the camera pose geometry in various earth regions. GE (Google Earth pro 7.3.4.8248, Accessed in 2022) is an application that provides satellite images for various regions of the earth. If the 3D data can be reconstructed using GE satellite images, it is possible for common users to obtain a 3D surface in various regions that are not reachable. If the data acquisition date is changed in GE, the satellite images from various view directions can be acquired in the same region. Therefore, we propose the 3D surface reconstruction scheme using only multi-view and multi-date GE satellite images. An important problem is that the RPC corresponding to the GE satellite image is unknown. Therefore, we assume that the GE satellite image is a

perspective projection model (PPM), which is a pinhole camera model in computer vision. Our scheme is as follows: The first step is that the camera pose geometry (camera intrinsic and extrinsic parameters) between the GE satellite images is estimated using PPM based SfM method. Second, 3D surface reconstruction is performed using the Enhanced Soft 3D reconstruction algorithm (EnSoft3D) (Lee et al., 2021 and Lee et al., 2022) algorithm, which is the MVS method proposed in our previous research. Importantly, GE satellite imagery is a pseudo-ortho image. The original satellite image is transformed for integration into GE raster image. Therefore, the low-rise structure and ground are transformed similarly as orthoimage. Instead, the high-rise structure (such as building, apartment, tower, etc.) is not correctly transformed (it still has facade regions). For this reason, the error is included in the camera pose geometry computation in the SfM process. Thus, the high-rise structures are slantly reconstructed because the 3D space is the distorted hexahedron space (the perspective camera frustum is slanted). The key to solving this problem is that the satellite 3D surface is a weak perspective projection model (weak PPM). Because the height of reconstructed 3D surface is significantly smaller than the distance between the camera (satellite) and the ground. The weak PPM can be expressed as the 3D cuboid space of the orthographic projection model. Based on this theory, the last step is to compute the 3D homography to transform between distorted hexahedron space and orthographic cuboid space. Then, the 3D homography-based projective reconstruction is performed about the slanted 3D surface.

The proposed method can reconstruct the 3D surface of various regions using only GE satellite images. Another advantage is that

* Corresponding author

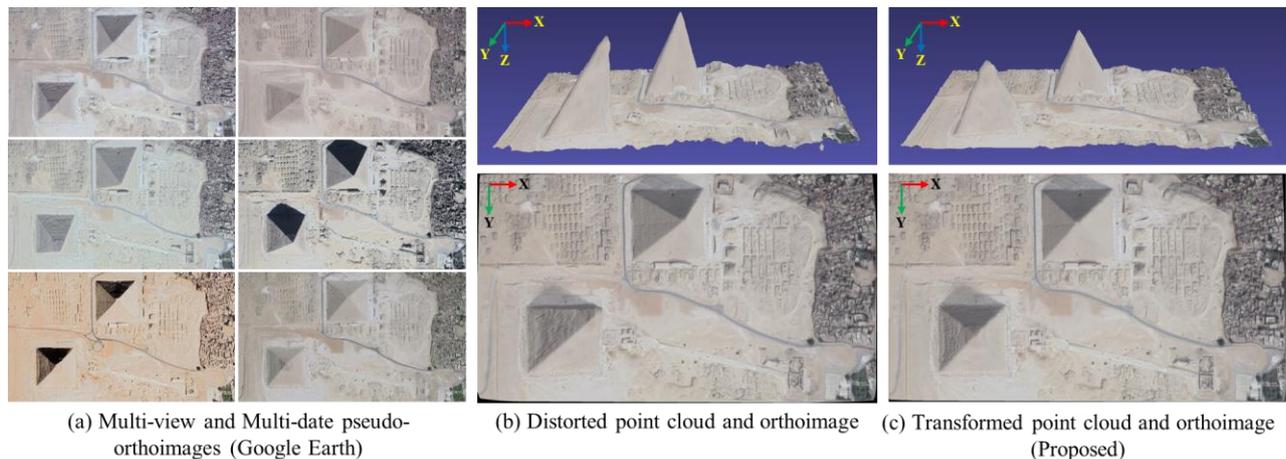


Figure 1. The 3D surface reconstruction result of Egypt-pyramid GE satellite images. (a) is multi-view and multi-date pseudo-orthoimages of Google Earth, (b) is the result of point cloud and orthoimage using MVS and incorrect camera pose geometry. The high-rise structure is reconstructed distorted (slantly). And (c) is the result of proposed method. The distorted point cloud is transformed as the orthographic projection model space using 3D homography-based projective reconstruction. Therefore, the correct orthoimage can also be generated.

orthoimage is simply generated using projection as height direction because the projective reconstructed 3D surface is orthographic projection model. Figure 1 shows the point cloud and orthoimage before and after projective reconstruction.

More details of the proposed method are described as follows: Section 3 explains the 3D surface reconstruction scheme of our EnSoft3D. And section 4 introduces the 3D homography-based projective reconstruction. The experimental results are shown in Section 5.

2. RELATED WROKS

Various MVS techniques are researched on a 3D surface reconstruction as multi-view satellite images. The most basic method is to use a multi-view satellite image and corresponding RPC. Facciolo (Facciolo et al., 2017) proposed the 3D surface reconstruction method using multi-view and multi-date satellite images and corresponding RPCs. If satellite images are input, the image pairs are selected. The stereo matching scheme is performed for 3D surface reconstruction using each pair. Lastly, the 3D surface of all pairs are aligned and fused for generating fused DSM.

Recently, the 3D surface reconstruction using PPM-based camera pose geometry has been researched. It has the advantage that various MVS or SfM methods of computer vision can be used. Zhang (Zhang et al., 2019) proposed the 3D surface reconstruction method (VisSat) of multi-view satellite images using SfM method. The PPM-based camera pose geometry is approximately computed using RPCs. And the 3D surface is reconstructed using COLMAP (Schönberger et al, 2016), which method is a high-performance SfM method among open sources. The author proves that the computer vision technique is competitive for 3D surface reconstruction of satellite images. On the other hand, Bullinger (Bullinger et al., 2021) uses various MVS methods, such as VisSat, MVE (Fuhrmann et al., 2014), COLMAP, OpenMVS (Cernea, 2020), for reconstructing satellite 3D surfaces. Instead of RPC, the PPM-based camera pose geometry is estimated using the SfM method. The skew information of the image sensor is not estimated in open-source SfM methods. Therefore, the author computes the skew parameter, so that the correctly 3D surface is reconstructed. In

various experiments, it proves that the PPM can reconstruct high-quality 3d surface of satellite images, not using RPCs.

However, as explained in the previous section, the GE satellite image is the pseudo-orthoimage. Therefore, a lot of error is included in the PPM-based camera pose geometry computed by

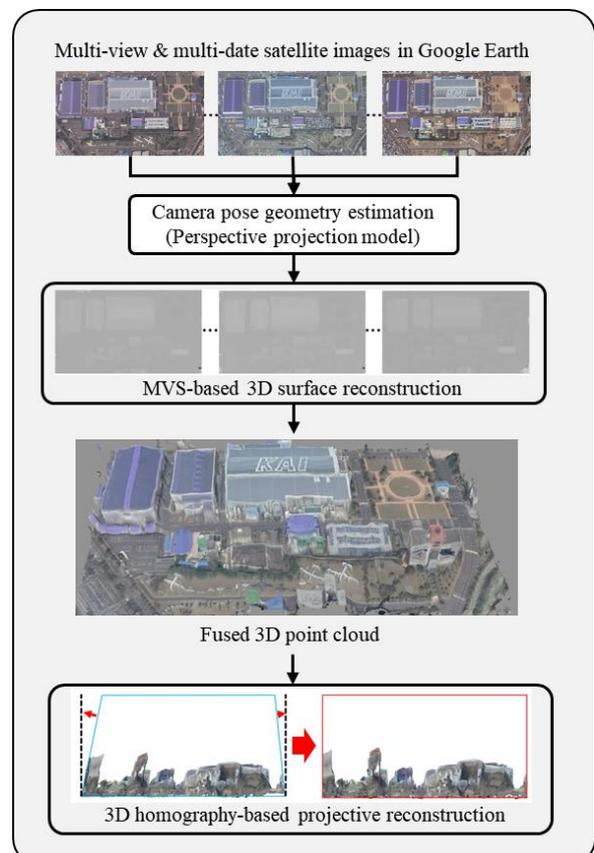


Figure 2. The proposed multi-view and multi-date satellite 3D reconstruction scheme.

SfM methods. This is the main cause for reconstructing distorted 3D surfaces. The proposed method solves this problem using a 3D space transformation using 3D homography-based projective reconstruction. This transformation calibrates the distorted 3D surface to orthographic projection model space. Therefore, the correct shape of the 3D surface can be obtained. The pipeline of the proposed method is shown in Figure 2.

3. 3D SURFACE RECONSTRUCTION SCHEME

In our previous research, the EnSoft3D uses the volumetric-based MVS method, Plane Sweep Stereo (Collins, 1996). Volumetric means that the matching cost volume is computed between the reference view and neighbor views. The volume is generated by 3D space discrete sampling using the image resolution plane (plane sweeping) and sampled 3D points are expressed as voxels. The matching costs (color similarity) of each voxel are computed by projection and back-projection based on camera pose geometry. It has the advantage that the matching costs are simply accessed. Another advantage is that the matching cost computation error can be refined by filtering or optimization, thus the accurate disparity can be obtained.

As similar as the Sebastian Bullinger method, we compute the PPM-based camera pose geometry (focal length, principal point, camera extrinsic parameter) using GE satellite images. It is performed using the image feature-based SfM method, COLMAP. If all camera pose geometry of multi-view and multi-date GE satellite images was computed, the matching cost volume is generated by plane sweep stereo method. The initial disparity map of each view is generated by WTA manner. And then, the iterative refinement method is performed to increase the accuracy of matching cost and disparity.

EnSoft3D generates the two volumes of each view, the surface consensus (consensus) and view soft visibility (visibility). The size of these two volumes is the same as the matching cost volume. Consensus is consist of the probability, which 3D point(voxel) is surface or not. If the textureless region in the image, the consensus is approximately computed by the plane information. Consensus is used to decrease the matching cost noise in the EnSoft3D iterations. Another volume, visibility, has the visible probability of which 3D point is occluded or not. The purpose of visibility is to refine the disparity of occlusion areas. These two volumes are computed by disparity maps of all views. The accuracy of disparity maps is increased by iterations. Since more accurate consensus and visibility are computed by refinement, the performance of matching cost update and occluded disparity refinement is increased by iteration. In our experiment, the number of iterations is 5 ~10 times.

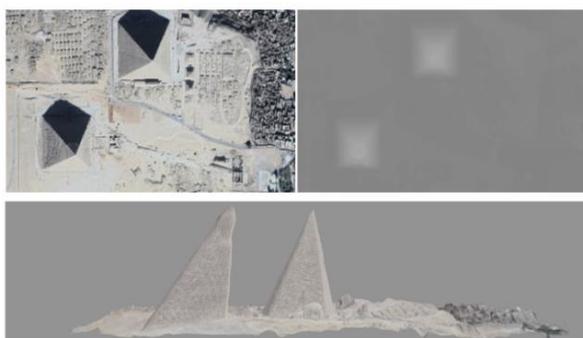


Figure 3. 3D reconstruction result of MVS (Top: input GE image and disparity map, bottom: point cloud). The point cloud shows that the high elevation structure is slanted.

EnSoft3D methods generate disparity maps of all views simultaneously. Therefore we fuse the point cloud of all views using the volumetric integration method, truncated signed distance field (Newcombe, 2011). Figure 3 shows the result of 3D surface reconstruction on Egypt-pyramid regions. Since the incorrect camera pose geometry is used, the pyramid is reconstructed slantly.

4. 3D HOMOGRAPHY-BASED PROJECTIVE RECONSTRUCTION

As described in the previous section, the GE satellite image is a pseudo-orthoimage. Figure 4 (top) shows the comparison of correct orthoimage and GE satellite image about the same regions. In GE, the high-rise structures have facade regions. However, the low-rise structures are similar to the correct orthoimage. This is the reason why the camera pose geometry is computed incorrectly. In Figure 4 (bottom), the 3D surface is reconstructed in distorted hexahedron spaces (green line).

In ideal, the height of PCM-based reconstructed 3D surface is much smaller than the distance between the ground and camera (satellite), thus it is a weak perspective projection model (weak PPM). Weak PPM can express as the cuboid space of orthographic projection model (orange line in Figure 4). Based on this theory, the proposed method transforms 3D surface from a distorted hexahedron space to cuboid space. The transformation between 3D spaces is the projective reconstruction. The 3D transformation is possible using 3D homography (4×4 matrix). It is computed using eight vertex pairs of two 3D spaces. The eight vertex pairs are selected as follows: The ground surface in the pseudo-orthoimage is transformed similar to the correct orthoimage. Therefore, the ground 3D points of distorted hexahedron space and orthographic cuboid space are shared. For this reason, the four vertices pair are selected from the 3D points on the ground surface. In the proposed method, the 3D points of the image corner are assumed with the ground surface and select them. Another four vertices in the distorted hexahedron space are

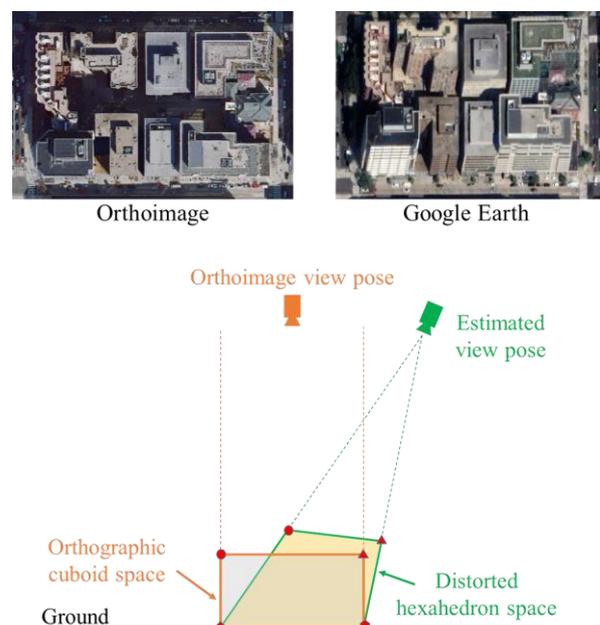


Figure 4. Comparison of reconstruction space between original orthoimage (orange) and GE (green) image in Washington, DC (DC GIS Group, 2021). The red points are the matching vertices of two spaces.

decided on the pixel-ray of the image corner. The corresponding four vertices in the orthographic cuboid space are decided on the perpendicular vector to the ground surface. These four vertex pairs of two spaces are selected with some interval, the interval is a user parameter. If all vertex pairs (V_{ppm}, V_{opm}) are selected, the 3D homography (H) is computed using singular value decomposition (SVD), as equation follows:

$$V_{ppm} = [v_{p,0}, \dots, v_{p,7}] \quad , \quad V_{opm} = [v_{o,0}, \dots, v_{o,7}] \quad (1)$$

$$H = SVD(V_{ppm}, V_{opm}) = \begin{bmatrix} h_1 & h_2 & h_3 & h_4 \\ h_5 & h_6 & h_7 & h_8 \\ h_9 & h_{10} & h_{11} & h_{12} \\ h_{13} & h_{14} & h_{15} & 1 \end{bmatrix} \quad (2)$$

$$P' = H \cdot \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix} \quad (3)$$

where $v_{p,0} \sim v_{p,7}$ = vertices of distorted hexahedron space
 $v_{o,0} \sim v_{o,7}$ = vertices of orthographic cuboid space
 p_x, p_y, p_z = distorted 3D point
 P' = transformed 3D point by 3D homography

The distorted 3D surface is equally slanted as the distorted hexahedron space. Therefore, the projective reconstruction is performed into orthographic cuboid space using 3D homography.

Because of 3D homography-based projective reconstruction, the correct shape of 3D surface can be obtained.

5. EXPERIMENTS

In the experiments, we obtained the GE satellite images from several regions such as South Korea, Egypt, and Bangkok. Figure 5 shows the result of 3D surface reconstruction and compared with COLMAP. Since the reconstruction of MVS method and point cloud fusion of all views, the dense point cloud can obtain. The facade region is also reconstructed because a lot of different views are used. Figure 5 (bottom) shows the height-based colorized point cloud (the blue is high, the red is low) of the proposed method and COLMAP. Since the inaccurate camera pose geometry is computed by COLMAP, the reconstructed 3D surface is distorted, and the depth scale is also inaccurate. Instead, the proposed method is correctly projective reconstructed. Another advantage is that the depth scale can be controlled by the interval of the user parameter. In our result of Egypt data, the height of the pyramid is correctly adjusted. Importantly, multi-date GE satellite images have a color difference by weather or data acquisition time, as shown in Figure 1 (left). However, the correct disparity can be estimated in MVS method by pixel matching of various views.

EnSoft3D can estimate a more dense 3D point cloud than COLMAP. It has the advantage in generating orthoimages. The orthoimage is simply generated by projection of height direction because the 3D surface is projective reconstructed into orthographic cuboid space. The result of the orthoimage is shown

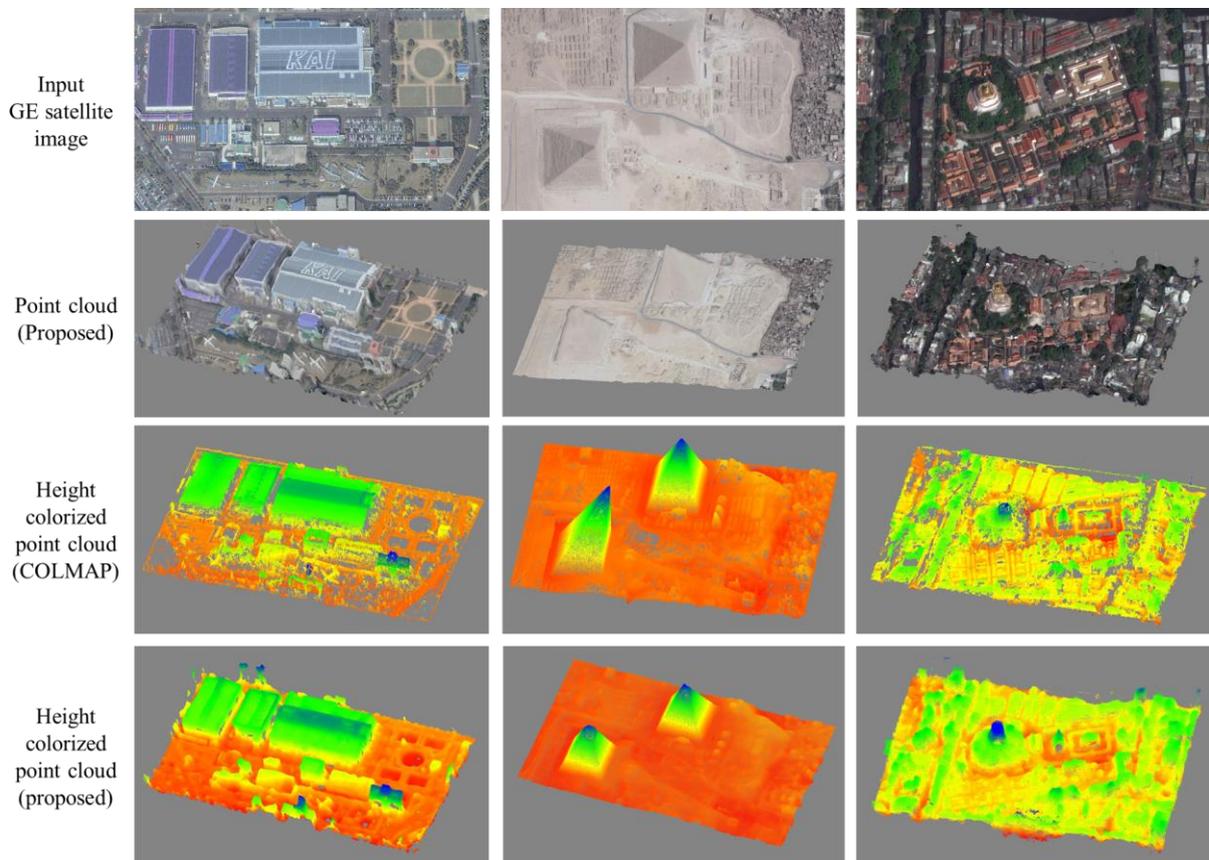


Figure 5. The result of 3D surface reconstruction using GE satellite images (South Korea, Egypt, Bangkok). The Height colorized point cloud is the colorization about the height value (the blue color is high, and the red color is low). The EnSoft3D can obtain more dense point cloud than COLMAP. It has an advantage for generating the orthoimage using projection.

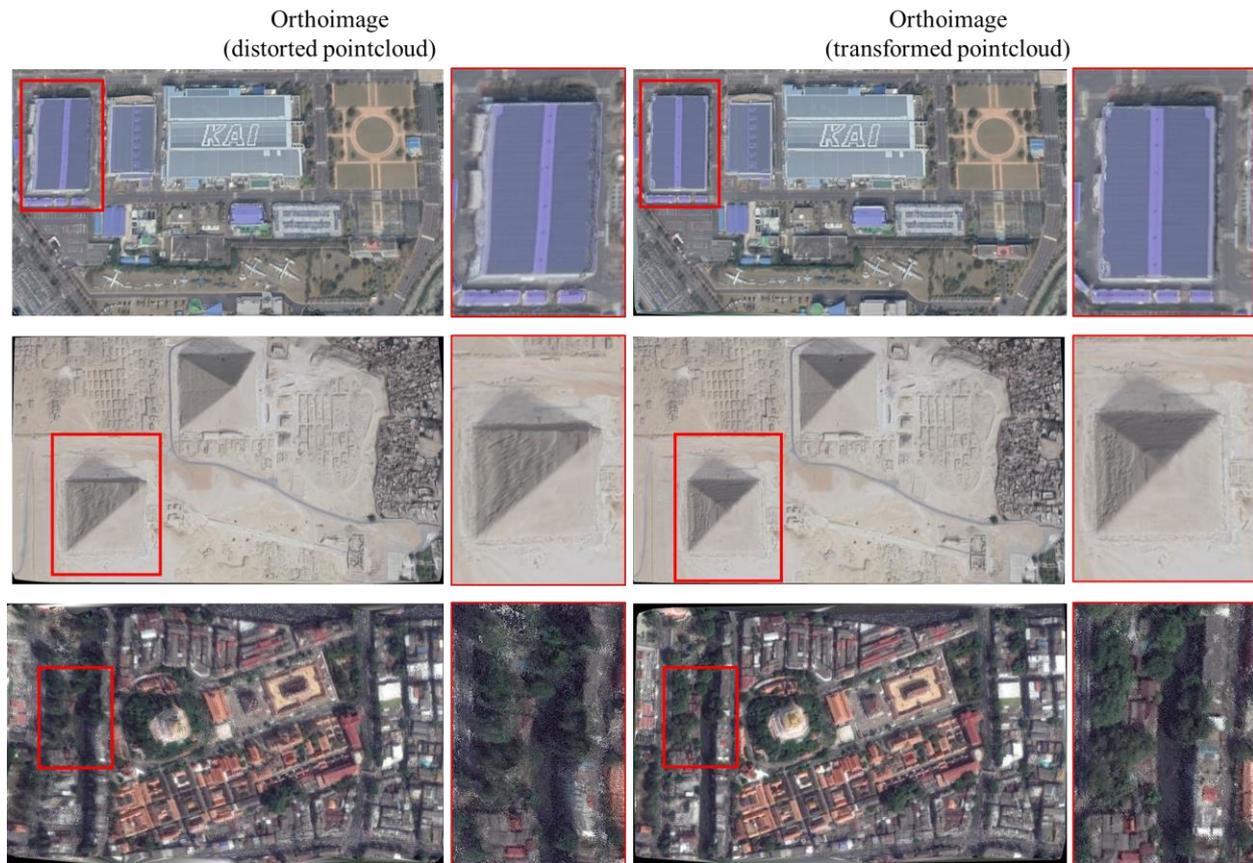


Figure 6. The comparison of generated orthoimage. Left images are the orthoimages using the distorted point cloud. It has still facade regions, similar to the GE satellite image. Instead, the right images show that more correctly orthoimage is generated (especially building, pyramid, and tower). It means that the 3D homography-based projective reconstruction can generate the correct shape of the 3D surface of an orthographic projection model space.

in Figure 6. In Figure 6(left), the distorted 3D surface (especially the building, pyramid, and tower) is generated similar to the GE satellite images in Figure 5(top). Instead, the projective reconstructed 3D surface can generate an accurate orthoimage. It is shown that the distorted 3D surface is correctly transformed into an orthographic projection model by 3D homography-based projective reconstruction.

6. CONCLUSION

In this paper, we proposed a 3D surface reconstruction scheme using GE satellite images. In GE, common users can simply acquire multi-view and multi-date satellite images. However, the RPCs of the GE satellite images are unknown. Therefore, we assume that the GE satellite image is the PPM of the pinhole camera model. And the SfM method is performed to compute the PPM-based camera pose geometry. After the camera intrinsic and extrinsic parameter is computed, the 3D surface is reconstructed by the MVS method. However, the GE satellite image is pseudo-orthoimage, not correctly transformed into orthoimage. This is the reason that the incorrect camera pose geometry is computed by SfM (slanted camera frustum). And the reconstructed 3D surface is distorted. The proposed method performs 3D homography-based projective reconstruction for solving this problem. The PPM-based satellite image is a weak PPM. The 3D space of weak PPM can express as an orthographic projection model. Therefore, we define eight vertex pairs about distorted

hexahedron space and orthographic cuboid space. And the 3D homography is computed using vertex pairs. The distorted point cloud can transform between 3D spaces by a projective reconstruction based on 3D homography.

Our MVS method can reconstruct dense 3D surfaces using GE satellite images. And the transformed 3D surface is reconstructed in the orthographic cuboid space. Therefore, the orthoimage can be generated simply using the projection of height projection. In the experiments, the orthoimages of the transformed 3D surface are more correctly generated than the distorted 3D surface. It shows that the 3D homography-based projective reconstruction can correctly solve the pseudo-orthoimage distortion problem. However, in the proposed method, the vertex pairs between two spaces are selected by the user parameter (interval). And the reconstructed 3D surface is not real size due to camera pose geometry. In future research, we will select the eight vertices automatically at the orthogonal cuboid space of real space. Due to this selection, it will obtain both the 3D surface of real space and solve the pseudo-orthoimage distortion problem simultaneously.

ACKNOWLEDGEMENTS

This work has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(No. 2021R1A2C2009722) and supported by Basic Science Research Program through the National Research Foundation of Korea

(NRF) funded by the Ministry of Education (No. 2021R1A6A1A03043144).

REFERENCES

- Hirschmuller, H., 2005: Accurate and efficient stereo processing by semi-global matching and mutual information,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, 807–814,
- Kendall, A., Martirosyan, H., Dasgupta, S., et al., 2017, End-to-end learning of geometry and context for deep stereo regression,” in *Proceedings of the IEEE international conference on computer vision*.
- Im, S., Jeon, H. G., Lin, S., and Kweon, I. S., 2019, DPSNet: end-to-end deep plane sweep stereo, in *7th International Conference on Learning Representations*.
- Xu, Q., and Tao, W., 2020, Pvsnet: pixelwise visibility-aware Multiview stereo network,” <http://arxiv.org/abs/2007.07714>.
- Google Earth Pro 7.3.4.8248. Accessed 2022. [online]: <https://earth.google.com/web/>.
- Lee, M. J., Um, G. M., Yun, J., Cheong, W. S., Park, S. Y. 2021: Enhanced Soft 3D Reconstruction Method with an Iterative Matching Cost Update Using Object Surface Consensus. *Sensors*, 21(19), 6680.
- Lee, M. J., & Park, S. Y. 2022: Refinement of Inverse Depth Plane in Textureless and Occluded Regions in a Multiview Stereo Matching Scheme. *Journal of Sensors*, 2022, 1 - 13.
- Facciolo, G., De Franchis, C., & Meinhardt-Llopis, E. 2017. Automatic 3D reconstruction from multi-date satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*.
- Zhang, K., Snavely, N., & Sun, J. 2019. Leveraging vision reconstruction pipelines for satellite imagery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*.
- Schönberger, J. L., Zheng, E., Frahm, J. M., Pollefeys, M. 2016. Pixelwise view selection for unstructured multi-view stereo. *Proceedings of the European Conference on Computer Vision (ECCV)*, Amsterdam, Netherland.
- Bullinger, S., Bodensteiner, C., & Arens, M. 2021. 3D Surface Reconstruction From Multi-Date Satellite Images. arXiv preprint arXiv:2102.02502.
- Fuhrmann, S., Langguth, F., Goesele, M., 2014. Mve: A Multiview reconstruction environment. *Eurographics Workshop on Graphics and Cultural Heritage*.
- Cernea, D., Accessed 2020. OpenMVS: Multi-view stereo reconstruction library. [online]: <https://github.com/cdseacave/openMVS>
- Collins, R. T., 1996. A space-sweep approach to true multi-image matching. *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*
- Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J.,... & Fitzgibbon, A. 2011. Kinectfusion: Real-time dense surface mapping and tracking. In *10th IEEE international symposium on mixed and augmented reality*.
- DC GIS Group, Accessed 2022. Aerial Photography Image Service (Orthophoto). opendata.dc.gov/datasets/DCGIS::aerial-photography-image-service-orthophoto-2021/explore (1 November 2021).