3D building reconstruction and validation using high-resolution stereo data

Jai Gopal Singla* and Sunanda Trivedi

Space Applications Centre, Indian Space Research Organisation, Ahmedabad 380 015, India

In recent times, due to advancement in satellite sensor technology and acquisition of very high resolution image data (0.3–1 m), high-resolution elevation models can be generated using satellite photogrammetry. High-resolution and precise digital elevation model (DEM) is an essential requirement for many applications like city modelling, terrain modelling, population estimation, resources estimation, visibility analysis, solar energy calculation, etc. Using high-resolution DEM and digital terrain model (DTM), heights of objects (buildings and trees) over the earth can be extracted for various applications. For visualization purpose, ortho image, building shapes with height information and important geographic information system information are used. This study mainly discusses the generation of high-resolution DEM from satellite stereo data, and processing of DEM and DTM. It also discusses extraction and validation of building heights using reference ground control points and visualization of highly accurate three-dimensional city model over the study area. This has been carried out using indigenous copyrighted software 'Generation of virtual 3D city model' of Space Applications Centre/ Indian Space Research Organisation.

Keywords: Building reconstruction, city modelling, satellite data, stereo processing.

VIRTUAL three-dimensional city models are used in fields like modelling of urban areas, disaster management, educational purposes, tourism and change detection. There are several applications built over city models, such as population estimation[1], navigation[2], CCTV camera deployment[3], visibility analysis[4], solar energy calculation[5] and many more. According to a study in Singapore, virtual 3D city models are employed in 29 use cases that are part of more than 100 applications[6]. At present many high-resolution sensors are in orbit, e.g. Cartosat 2S, Cartosat-3, Ikonos, WorldView-2 and WorldView-3. These sensors have the capability to acquire along-track/across-track stereo imagery. To generate digital elevation models (DEMs) from satellite-based optical stereo datasets, the popular approach of satellite photogrammetry is used. Aerial methods such as aerial photogrammetry using drones/unmanned aerial vehicles (UAVs) and light detection and ranging (LiDAR) scanning methods are relatively costly and require a lot of time in planning and security clearances. But DEMs and ortho imagery generated through drones using LiDAR and high-resolution cameras are relatively more accurate. Satellite photogrammetry has advantages of very wide coverage in single pass and is cost-effective. In recent times, due to advancements in satellite sensor technology and acquisition of very high resolution (VHR) image data (0.3–1 m), high-resolution elevation models can be generated.

In this study, raw stereo images from WorldView-2 satellite are used to extract building height. Stereo images are processed using satellite photogrammetry and high-resolution DEM is generated over the study area. DEM represents the surface along with object height, whereas digital terrain model (DTM) represents only the bare earth surface. Objects such as buildings, bridges and trees are filtered from DEM and high-resolution DTM is generated. Orthorectified image is generated using DEM and sensor parameters as input. Building height is extracted using DEM, DTM and building shape files. Further, the extracted building height is validated with ground control points (GCPs) over the study area to assess efficacy of the result. Finally, virtual 3D city modelling is carried out using an in-house developed software and a novel approach as mentioned in our earlier work[7]. Accuracy of the rendered virtual 3D city model primarily depends on accuracy of the generated DEM, DTM and ortho imagery.

Across the world, many academicians and researchers have worked on this interesting problem of generation of high-resolution DEM using high-resolution stereo pairs for application in diverse fields. In 2002, Fraser *et al.*[8] reconstructed a building from the processing of 1 m stereo Ikonos imagery over a 7 sq. km small area at the University of Melbourne, Australia. Liu *et al.*[9] assessed the quality of building height extraction from ZiYuan-3 multi-view imagery and WorldView-2 images. They have evaluated 380 reference buildings and achieved a RMSE of 13.08 m. Ghuffar[10] used multi-date and multi-view PlanetScope imagery and achieved normalized median of absolute deviation (NMAD) of elevation difference between computed DEM and LiDAR DEM as 4.1 m. Gong *et al.*[11] performed a detailed study about the generation of digital surface model using high-resolution satellite stereo imagery of WorldView-2 sensor. They have generated a dense DEM using dense image matching algorithm of the

*For correspondence. (e-mail: jaisingla@sac.isro.gov.in)

LibTsgm library and achieved median height difference of <1 m between generated and reference DEM. Wang *et al.*[12] studied DEM generation from WorldView-2 stereo imagery for its application in active tectonics. Their findings indicate that the DEM generated from Worldview-2 stereo imagery is capable of measuring relative deformed topographic features. Nemmaoui *et al.*[13] studied the generation of DEM and DTM using WorldView-2 stereo images for applications in agriculture under plastic covered greenhouses (PCGs). They performed 3D mapping over PCG areas and validated the results with LiDAR data, with a mean error of <0.5 m. Aguilar *et al.*[14] assessed the quality of DEM generated from WorldView-2 and WorldView-3 sensors over different land covers. They concluded that DEM accuracies are slightly better in the case of WorldView-3 compared to WorldView-2.

In this study, high-resolution DEM is generated using the concepts of satellite photogrammetry. High-resolution DTM is generated by filtering DEM using slope filters[15] and ortho image (nadir view) is generated using high-resolution DEM and sensor parameters. Building height information is computed using DEM, DTM and building shape. The building height is validated with respect to GCPs and high-quality virtual 3D city model is visualized over the city area. The primary objective of this study is to achieve absolute accuracy of 2 m in the vertical direction, which can be useful in identifying floor-level changes in buildings. Another objective of this study is to provide a visual analysis platform to identify the small/mid-level and high-rise buildings in the city area.

Study area

Ahmedabad city is one of the largest cities and former capital of the Indian state of Gujarat. It consists of all types of building, ranging from low-rise to mid-level and high-rise. It is one of the most populous cities in India. Due to the availability of all types (size and shape) of building structures, as well as raw stereo data and reference ground data, we have chosen Ahmedabad as the study area.

Dataset

WorldView-2 sensor provides panchromatic imagery with a resolution of 0.46 m and multispectral band image with a resolution of 1.85 m (ref. 16). The Digital Globe team supplied the input data to evaluate the applications in accurate city modelling[17]. A detailed description of the datasets is given below.

(a) High-resolution imagery: Multi-spectral high-resolution image (<2 m) from WorldView-2 sensor.
(b) Stereo data: A pair of in-track stereo images of 50 cm resolution was acquired using WorldView-2 sensor in 2015 over the study area. It has in-track view angles of 24.5° and –4.2° respectively. Images have good overlap (99.86%) and base to height ratio (*B*/*H*) ratio (0.53) to compute high-resolution DEM. Table 1 summarizes the important parameters of the stereo pair.
(c) OpenStreetMap files: OpenStreetMap (OSM) has collected important data on all countries across the globe over a decade using GPS surveying and crowd sourcing. Different vector layers are available with information about two-dimensional buildings, important landmarks, road networks, waterways information and administrative boundaries. OSM data over the study area (~5 m spatial resolution) for 2014–15 were downloaded free of cost from the GeoFabrik website[18].
(d) Ground control points: Well-distributed GCPs generated with GPS dual frequency receivers and processed with precise ephemeris data were used for the generation of DEM and evaluation purposes. These GCPs have an accuracy <1 m.

Methodology

The proposed methodology covers important steps utilized for the generation of high-resolution DEM using stereo images, extraction of high-resolution DTM, generation of ortho-rectified imagery and extraction of height of each building shape. Generation of the virtual 3D city model is also explained. The following steps are used for DEM generation.

Identification of ground control points

GCPs are permanent features on the earth's surface which can be identified on satellite images precisely and have known *X*, *Y* and *Z* (elevation) ground coordinates. GCPs can be collected using large-scale maps, ground surveys, ground GPS observations, orthorectified images and DEMs. To model image and ground space, a minimum of GCPs is required. However, to increase accuracy (using least square solution), more GCPs are recommended. Ten well-distributed GCPs were identified in the overlap

Table 1. Characteristics of panchromatic images from WorldView-2

Images	WorldView-2	
	Image 1	Image 2
Acquisition date (D/M/Y)	2015-01-10	2015-01-10
Acquisition time (GTM)	06:06	06:07
Scan direction	Forward	Forward
Off-nadir view angle (°)	25.1	8.4
In-track view angle (°)	24.5	–4.3
Cross-track view angle (°)	–5.3	–7.2
Collected GSD (m)	0.58	0.47
Sun azimuth (°)	157.7	158
Sun elevation (°)	41.9	42

region with respect to the reference points to obtain accurate positional and rotational coefficients.

Tie point/match point generation using image matching technique

A tie point/match point is a similar feature precisely identified in the stereo pair images. ERDAS Imagine provides the facility to generate automatic match points based on image matching techniques such as area-based matching using correlation, cross-correlation and other feature-based techniques[19]. More than 9000 tie points were identified in the stereo pairs, which helped generate a high-resolution and accurate DEM.

Block triangulation

This is a process which defines the mathematical relationship between the images, sensor model and the ground[20]. If this mathematical relationship is defined accurately, it will result in better accuracies of the processed image. Block triangulation was carried out with RMSE accuracy of one-tenth of the pixel.

DEM interpolation/surface fitting

In the domain of image processing, interpolation refers to determining the value of a point using known values of neighbourhood points. There can be holes in DEM due to various reasons like shadow effects, non-availability of matching feature points in both images, etc. Points obtained by image matching are a set of irregular points that do not completely represent the surface[21]. Therefore, regular DEM was generated at 1 m grid interval using block triangulation and DEM interpolation.

Orthorectification

This process corrects a raw image for geometrical errors and displacements caused by the undulated terrain[22]. An orthorectified (near-nadir view) image was generated over the study area using one of the raw images from the stereo pair, sensor parameters and DEM.

Visual analysis of high-resolution DEM

High-resolution DEM with 1 m resolution and orthorectified image were obtained using the above mentioned steps. The generated DEM and ortho image were evaluated for their visual and geometric qualities with respect to GCPs. Here, DEM is classified according to height information for visual analysis. Figure 1 depicts the high-resolution DEM obtained after processing of stereo data. Values less than or equal to zero specify 0 m surface height, 0–7 m represents low-range height, 10–17 m mid-range height and >17 m represents high-range height. DEM was labelled/classified according to the height information. Figure 1 also shows the difference in building height over different areas.

Extraction of building height using DEM

QGIS is an open-source software GIS package used for raster and vector data processing and analysis[23]. In slope filtering algorithm, the slope or height difference between two nearby points is measured. If the slope is above a certain threshold, then the highest point is assumed to belong to an object[24]. The generated DEM was loaded in QGIS tool as a raster layer. DTM slope filter tool with search radius of 20 and terrain slope of 3 m (single-floor-level height) was used to appropriately filter the DEM into two parts: bare earth surface and image, including the objects (buildings, trees and other objects) over the earth's surface. Bare earth raster data (after removal of objects), which were the output of the DTM slope filter, were stored on the disk for further use. Bare earth raster data contained holes due to slope filtering. So, they were reloaded and interpolated using r.fillnulls tool[25] and spline

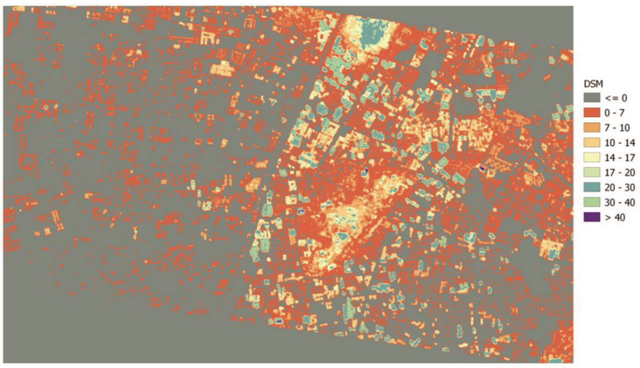

Figure 1. Digital elevation model (DEM) results and building height (m) classification using DEM data.

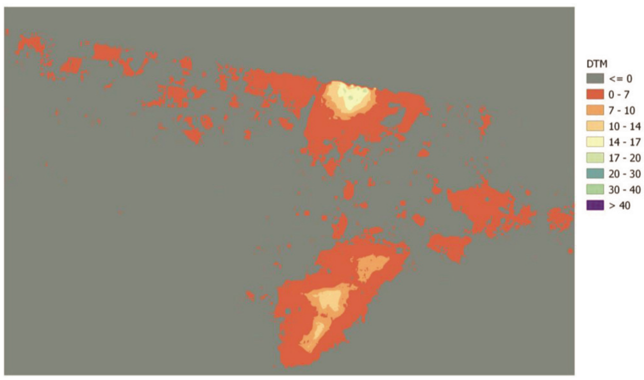

Figure 2. Digital terrain model (DTM) height (m) results after slope filtering.

Figure 3. DEM and DTM profile over selected buildings.

interpolation to fill the holes in the data (DTM) (Figure 2). Object heights were extracted by calculating the difference of input DEM layer and processed bare earth layer, i.e. DTM in principle. For extracting height of each building from raster datasets, Zonal Statistics tool[26], DEM, DTM data and building vector 2D shape files were used. All the required inputs were provided to the Zonal Statistics tool and the mean statistics calculated for each building shape. Mean values were chosen as the Zonal Statistics value to better represent height of that particular building. Using the Zonal Statistics module, DEM mean height and DTM mean height over each building footprint were calculated and appended with the attribute ID of each shape. Difference between the mean values of DEM and DTM was calculated to obtain the exact height of the building structure and stored in the table as a new attribute, namely Building_ht. The processed zonal stats table was combined with the original 2D shape file of the building to obtain the new output building shape file which contains the Building_ht attribute of each building against its identification ID and name.

Virtual 3D city model using indigenous software package

In order to generate the 3D city model correctly, pixel-level registration is mandatory. Sub-pixel level registration accuracies were obtained between orthorectified imagery and building footprints using the approach mentioned in our earlier work[7].

It is challenging to render millions of polygons in computer graphics along with raster and several vector layers over the study region. Yldrm *et al.*[27] developed a memory management approach for hydrological terrain processing. For real-time rendering of millions of polygons and virtual 3D city models, we have used OpenSceneGraph (OSG) library[28] and C++ language as programming interfaces and developed an indigenous copyrighted software known as 'Generation of virtual 3D city model'. This software takes orthorectified image layer, building shape files and other important GIS vector layers as input and provides a highly accurate and informative virtual 3D city model on the computer screen. The developed software has built-in features to support optimal rendering of millions of polygons, extrude height of building shapes and overlay textures over the buildings shapes. The software works well on Windows and Linux operating systems. It has been developed considering tilling/pyramid approach for memory and resource optimization using OSG[28] application programming interfaces (APIs). The tiling mechanism generates smaller tiles of image textures containing different levels of information with a typical tile size of 16 k × 16 k, and stores them in the main memory. Level-Of-Details mechanism implemented in OSG helps replace a tile by any one of the four sub-tiles depending on the viewer's position, and a high-resolution tile is rendered based on the range of the viewing coordinates[29].

Results and discussion

Initially, slope filtering over high-resolution DEM was evaluated. For validation of slope-filtering results, DEM and DTM values at various random locations were compared (Figures 3 and 4). It can be inferred from Figures 3 and 4 that the slope filter has correctly removed the height of the objects over the earth's surface and the surface with holes is also interpolated correctly in the DTM layer. Further, extracted heights over 55–56 random size buildings were validated with respect to known GCPs (Figure 5). Maximum error of 4.5 m and RMSE error of ~2 m were calculated in the vertical direction over these buildings. The height of 10 different building classes was also compared with ground truth. Table 2 show the results. Finally, the virtual city model was generated using this indigenous software developed at SAC/ISRO.

Figure 4. DEM and DTM profile over wider building coverage.

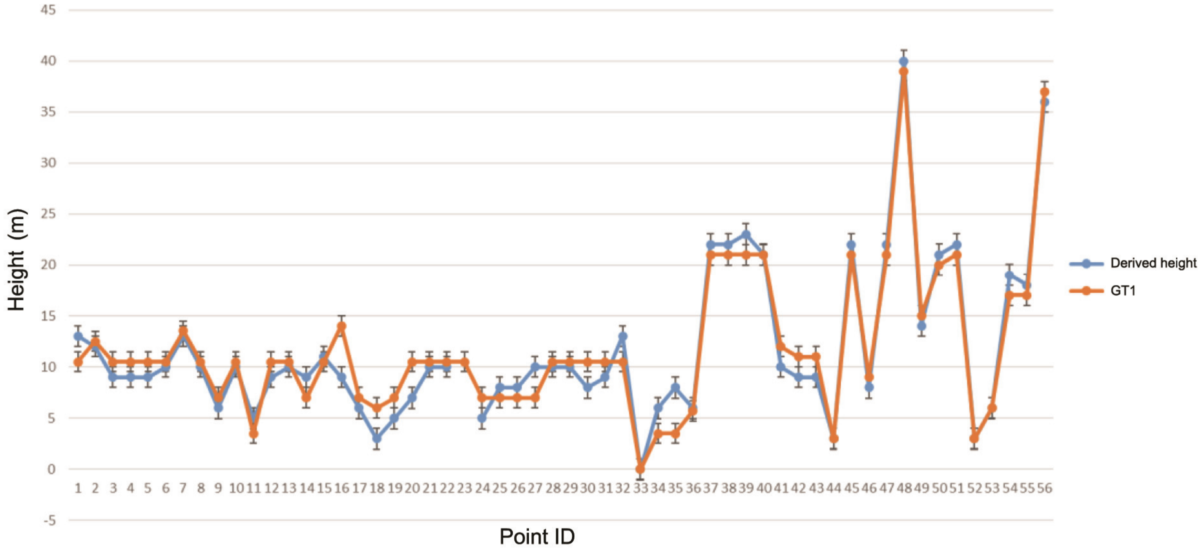

Figure 5. Extracted height versus ground truth over more than 50 buildings.

Table 2. Comparison of height between extracted versus ground data in different building categories

Extracted building height (m)	Ground data height (m)	Remarks
22	21	High-rise apartments
23	21	High-rise apartments
10	12	Mid-rise apartments
3	3	Dispensary
40	39	Hotel
14	15	Shopping mall
21	20	Shopping mall
18	17	Shopping complex
3	3	Temple
36	37	High-rise apartments

DEM and DTM result profiles

Using the methodology mentioned earlier in the text, DEM and DTM were generated over the study area. The DEM and DTM results were visualized (Figures 1 and 2 respectively) and it was interpreted that the slope filter had filtered out well the buildings and other objects (trees and bridges, etc.) from DEM data. In Figure 2, most of the data lie in 0–7 m range, which shows the height of the barren land at known areas in Ahmedabad. Further, slope filtering using DEM and DTM profile over multiple building locations in the city was validated. Figures 3 and 4 depict the DEM and DTM profile over building height in Ahmedabad. It can be observed that the slope filtered has efficiently removed the building structures and the surface has been smoothened using the interpolated values.

Accuracy and validation of building height

Accuracy of the 3D virtual city model depends on the accuracy of various input layers, and mainly on DEM and DTM accuracies. In this study, we have generated DEM at 1 m grid interval and further derived DTM using slope filter. Ground data were collected over more than 55 random and stable building structures using ground surveys,

Figure 6. Three-dimensional city model using building height over parts (northwest) of Ahmedabad city, Gujarat, India, captured using indigenous software.

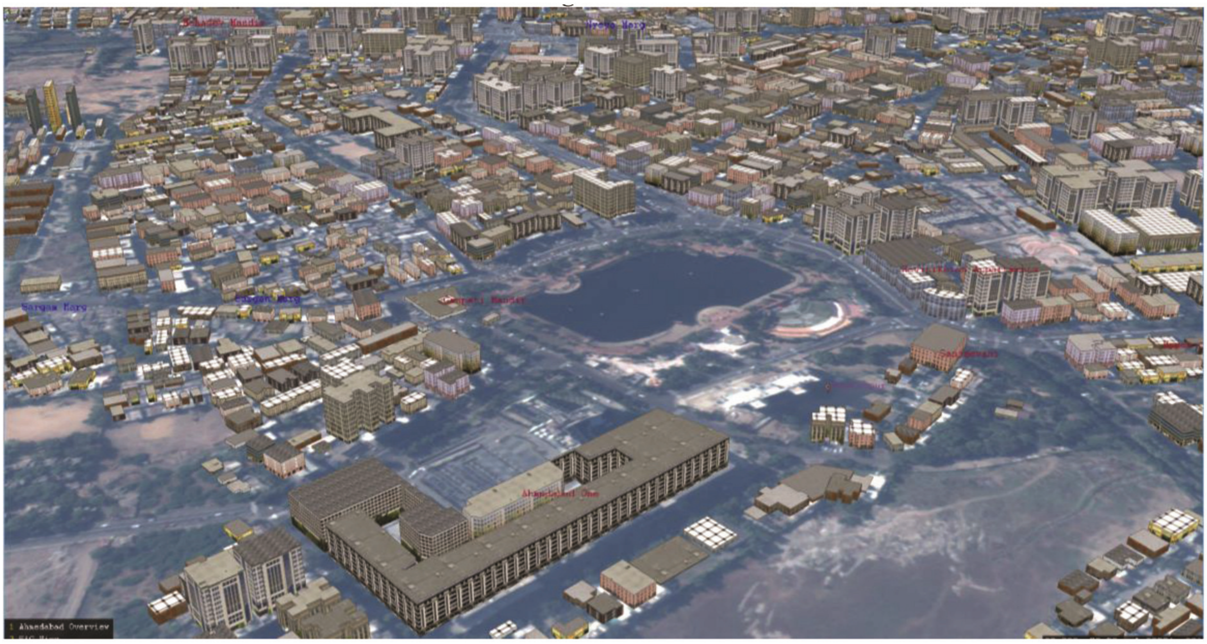

Figure 7. Three-dimensional city model over parts (northwest) of Ahmedabad city captured using indigenous software.

GPS and laser disto meter device. Vertical accuracy of the extracted building heights was assessed with respect to GCPs and RMSE of 2.2 m was obtained in the vertical direction, whereas planimetric accuracy was within 1 m. Maximum difference of 4.5 m between reference and measured building heights was noted. A comparison between reference height and extracted height of different building classes was made according to Table 2. As observed from Figure 5 and Table 2, the values of extracted height are close to the ground truth values. Also, the results are sufficient to identify small, mid-level (3–4 storeys) and high-rise buildings.

Visualization of results of city modelling

The virtual city model over the entire study area was generated using the indigenous software. The results help identify buildings of different shapes and sizes, heights and textures according to the extracted height information. Using features of the developed software, users can have different viewpoints of the scene depending on the view angle and altitude. Figures 6 and 7 show the 3D virtual city model over different areas of Ahmedabad. Buildings of different shapes, sizes and heights are shown in the figures.

Conclusion

In this study, satellite photogrammetry has been used to extract building height. Initially, a high-resolution DEM of 1 m was generated using a high-resolution satellite stereo pair. Slope filtering and interpolation algorithms were applied on DEM and high-resolution DTM was produced. Orthorectified (nadir-view) image was also generated using the raw image, sensor parameters and high-resolution DEM. Height of each building shape was extracted by taking inputs of DEM and DTM along with building shape. In order to assess the building height accuracies, height information over random and stable building structures was validated with the reference data. RMSE of ~2 m was obtained over 55 random buildings, which indicates detection of floor-level height in various types of buildings. Hence it is possible to determine the floor-level accuracies in building height using in-track/across-track satellite data. Also, virtual 3D city models are useful in applications such as smart city planning, population estimation, navigational aid, CCTV deployment, visibility analysis and solar calculators.

1. Guo, H., Cao, K. and Wang, P., Population estimation in Singapore based on remote sensing and open data, ISPRS Geospatial Week, 2017.
2. Cappelle, C., Najjar, M., Charpillet, F. and Pomorski, D., Virtual 3D city model for navigation in urban areas. *J. Intell. Robot Syst.*, 2011.
3. Yaagoubi, R., Yarmani, M., Kamel, A. and Khemiri, W., HybVOR: A Voronoi-based 3D GIS approach for camera surveillance network placement. *ISPRS Int. J. Geo-Inf.*, 2015.
4. Lonergan, C. and Hedley, N., Unpacking isovists: a framework for 3D spatial visibility analysis. 2015.
5. Liang, J., Gong, J., Zhou, J., Ibrahim, A. and Li, M., An open-source 3D solar radiation model integrated with a 3D geographic information system. 2014.
6. Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S. and Çoltekin, A., Applications of 3D city models: state of the art review. 2015.
7. Singla, J. and Padia, K., A novel approach for generation and visualization of virtual 3D city model using open source libraries. *J. Indian Soc. Remote Sensing*, 2020.
8. Fraser, C., Baltsavias, E. and Gruen, A., Processing of Ikonos imagery for submetre 3D positioning and building extraction. *ISPRS J. Photogramm. Remote Sensing*, 2002, **56**, 177–194.
9. Liu, C., Huang, X., Wen, D., Chen, H. and Gong, J., Assessing the quality of building height extraction from ZiYuan-3 multi-view imagery. *Remote Sensing Lett.*, 2017, **8**(9), 907–916; doi:10.1080/2150704X.2017.1335904.
10. Ghuffar, S., DEM generation from multi satellite PlanetScope imagery. *Remote Sensing*, 2018, **10**, 1462; doi:10.3390/rs10091462.
11. Gong, K. and Fritsch, D., A detailed study about digital surface model generation using high resolution satellite stereo imagery. *ISPRS Ann. Photogramm.*, 2016.
12. Wang, S. *et al.*, DEM generation from WorldView-2 stereo imagery and vertical accuracy assessment for its application in active tectonics. *Geomorphology*, 2019, **336**, 107–118; https://doi.org/10.1016/j.geomorph.2019.03.016.
13. Nemmaoui, A., Aguilar, F., Aguilar, M. and Qin, R., DSM and DTM generation from VHR satellite stereo imagery over plastic covered greenhouse areas. *Comput. Electron. Agric.*, 2019.
14. Aguilar, M., Nemmaoui, A., Aguilar, F. and Qin, R., Quality assessment of digital surface models extracted from WorldView-2 and WorldView-3 stereo pairs over different land covers. *GISci. Remote Sensing*, 2018.
15. Wichmann, V., DTM filter module, 2010; www.saga-gis.org/saga_tool_doc/2.2.5/grid_filter_7.html (accessed on 30 April 2021).
16. https://www.euspaceimaging.com/about/satellites/worldview-2 (accessed on 20 May 2021).
17. https://discover.digitalglobe.com (accessed on 30 April 2021).
18. https://www.geofabrik.de/geofabrik/openstreetmap.html (accessed on 30 April 2021).
19. https://www.hexagongeospatial.com (accessed on 20 May 2021).
20. Mikhail, E., Bethel, J. and McGlone, J., *Introduction to Modern Photogrammetry*, American Society of Photogrammetry and Remote Sensing (ASPRS), Louisiana, 2001, pp. 118–119.
21. Schenk, A., Automatic generation of DEM's. *Digit. Photogramm.: ASPRS*, 1996.
22. Mikuni, A., Digital ortho-photos: production, mosaicking and hard copy. *Digit. Photogramm.: ASPRS*, 1996.
23. www.qgis.org (accessed on 30 April 2021).
24. Sithole, G. and Vosselman, G., Report: ISPRS comparison of filters, Department of Geodesy, Delft University of Technology, The Netherlands, Commission III, Working Group 3, September 2003.
25. https://grass.osgeo.org/grass90/manual/r.fillnulls.html (accessed on 20 May 2021).
26. www.docs.qgis.org/2.18/en/docs/user_manual/plugins/plugins_zonal_statistics.html (accessed on 30 April 2021).
27. Yldrm, A., Watson, D., Tarboton, D. and Wallace, R., A virtual tile approach to raster-based calculations of large digital elevation models in a shared-memory system. *Comput. Geosci.*, 2015, **82**(C), 78–99; https://doi.org/10.1016/j.cageo.2015.05.014.
28. www.openscenepgraph.org (accessed on 20 May 2021).
29. Wang, R. and Qian, X., *Open Scene Graph 3 Cookbook*, 2012, chapter 3, pp. 119–120; ISBN 978-1-84951-688-4.

ACKNOWLEDGEMENTS. We thank Shri N. M. Desai, Director, Space Applications Centre for guiding and permitting the publication of this paper. Suggestions from internal referees to improve an earlier version of this paper are sincerely acknowledged.

Received 1 February 2022; accepted 10 February 2022

doi: 10.18520/cs/v122/i8/900-906