Scholars Journal of Engineering and Technology
OPEN ACCESS
Abbreviated Key Title: Sch J Eng Tech
ISSN 2347-9523 (Print) | ISSN 2321-435X (Online)
Journal homepage: https://saspublishers.com/journal/sjet/home

Virtual Reality Urban Modeling from Hybrid Image-based Sensors

Riyam H. Hadi[1*], Fanar M. Abed[2], Mustafa T. Mustafa[1]

[1]Engineering Technical College, Middle Technical University Baghdad Iraq
[2]College of Engineering, University of Baghdad Iraq

***Corresponding author:** Riyam H. Hadi
| **Received:** 10.05.2019 | **Accepted:** 17.05.2019 | **Published:** 30.05.2019
DOI: 10.36347/sjet.2019.v07i05.004

Abstract | **Review Article**

There is an increasing interest in generating virtual reality invironments due to the growing 3D visual media in the market. Aerial photogrammetric data is usually used to generate virtual urban models but there are some issues in creating 3D city model from aerial photos. This includes problems in identifying small buildings and limited building geometry detail levels (LOD) especially for building facades. To progress aerial photogrammetric data for the development of virtual city modeling, this research is carried out by integrating multi-source image-based data to verify LOD3 level of buildings accomplished by Close Range Photogrammetry (CRP). This study includes the integration of aerial and CRP images those delivered from Unmanned Aerial Vehicle (UAV) platform and hand-held static camera, for the construction of augmented reality model of the University of Baghdad (UOB) campus. Two buildings with different specifications inside the university campus are selected to apply the study and discuss results. In this study, the integration of multi-source image-based data, are achievable using different algorithms from commercial software. The results approved that using UAV platform to capture close-range images of very high building facades is recommended and can deliver superior results as compared with results delivered from static ground stations using hand-held cameras. The validation accuracy of the final virtual reality UOB model deliver RMSE of 0.271051m., 0.250703 m. and 0.301240 m. in easting, northing, and elevation respectively, which consider promising for small scale projects with small detailed objects in urban areas.
Keywords: Virtual Reality Urban Modeling, Hybrid Image-based.

INTRODUCTION

Data integration and registration means combination data from different sources and providing users with a unified data display [1]. It is usually involving the combination of multidisciplinary knowledge from multi-sources to produce high data quality [2]. Integrated data from multi-sources give better and more accurate products and facilitate processing data task [3]. Currently, the most effective techniques for surveying engineering applications and many other applications are photogrammetry in first place in addition to laser scanning [4]. Photogrammetry is not newly created technique, but has at least one century of trajectory and has evolved to reduce their costs as they have been advancing information technologies with the development of the photograph. On the other hand, laser scanning was applied and adopted in later years in different applications due to its superior accuracy and less user and data processing efforts in comparison with photogrammetry. However, it has an important side effect in this manner, which is high costly sensor. This makes laser scanning preferable as a second place technique after photogrammetry in some particular applications related to 3D object reconstruction and thus make it beyond the focus in this research.

Photogrammetry and imaged-based data give the possibility of acquiring an extraordinary high quantity and quality of data and dimensions, whose density is conditioned by the density of the camera sensor resolution. This is particularly depending on the distance at which the objects lie to the sensor and thus the sensor position whether it's aerial, terrestrial, or even carried by a far distance platform such as satellites [5]. The 3D application can be used both in small elementary objects, such as in the context of an excavation or structures, or more extensively, for the landscape and territory. Photogrammetry from terrestrial platforms has its limitations in terms of visibility of interested objects and thus the quality of the obtained results [5]. However, when it comes to obtaining data from inaccessible areas like roof of a building or some archaeological ruins, it requires delivering images from different platform cameras and later integrate these data to deliver the product [1]. This can be achievable by capturing images from terrestrial cameras and images from elevated cameras lies on aerial and satellite platforms, which can be easily delivered today from UAV cameras when low or even high altitude is required with high ground sampling resolution [6]. Every day employment is more widespread of aerial photography in any field, like the agricultural, agronomic, archaeological, architectural, military, in the reports videography of any activity, advertising, detection of fires and mountain rescues, etc., through the utilize of digital or thermal cameras [5]. Therefore, the combination of multi-source photogrammetric datasets (e.g. aerial, terrestrial, and satellite) is the ideal low-budget solution for obtaining global geometric database for any structure, as it obtains information about concealed or inaccessible sites or objects. For this reason, the main focus of this research was stated to deal with the integration of multi-source image-based datasets and highlighting the powerful of such integration to improve 3D modeling in virtual reality planning and design applications particularly.

Theoretical Concepts

As all professionals know, photogrammetry is representing the process of measuring, recording and interpreting photographs, electromagnetic radiation patterns and furthermore other phenomena [7]. Registration is an essential task in image processing utilized to match two or more images captured, for instance, at various times, sensors, or viewpoints [8]. The most popular transformation model utilized for handling image registration is 3D affine transformations. 3D affine transformations have been exceedingly utilized in computer vision, in special, in the field of model-based object recognition and may have included a various number of related parameters as shown in Figure (1) [9].

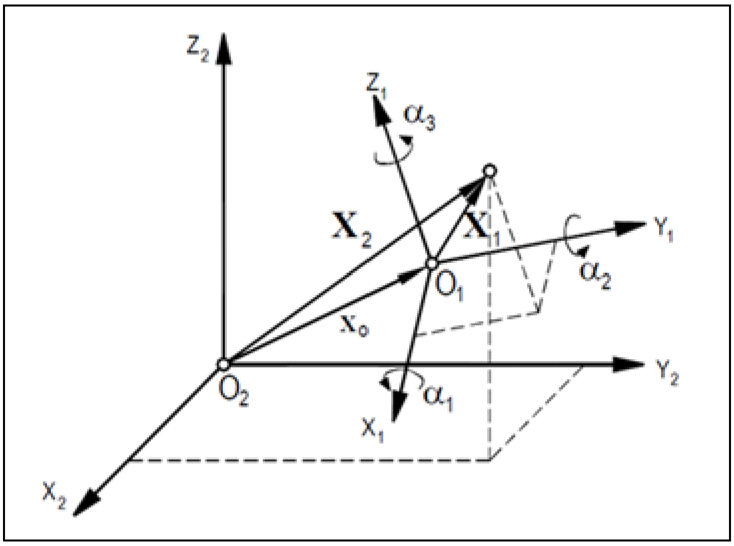

Fig-1: 3D Affine transformation concept [10]

With 9 parameters affine function (Helmert's transformation), for example, two sets of 3D Cartesian coordinates are considered (Figure (1)). Helmert's transformation can be formulated between these two sets of data as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{(2)} = \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} + \mu \cdot R(\alpha_1, \alpha_2, \alpha_3) \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{(1)} \qquad (1)$$

For simplicity, the above relationship can also be written as [11]:

$$\mathbf{X}_2 = \Delta\,\mathbf{x} + \mu \cdot \mathbf{R} \cdot \mathbf{X}_1 \qquad (2)$$

Where:
X1, X2: The vectors location of the same point, both in a fixed and transformed coordinate system.
Δx: the translation vector.
μ: the scale factor.
R: the rotation matrix.

The components of the translation vector are:

$$\Delta \mathbf{x} = \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} \quad (3)$$

Thus, the translation from the origin X1 into the origin X2 can be computed.
The rotation matrix is an orthogonal matrix consisting of three successive rotations and looks as follows:

$$\mathbf{R} = \mathbf{R}_3(\alpha_3) \cdot \mathbf{R}_2(\alpha_2) \cdot \mathbf{R}_1(\alpha_1) \quad (4)$$

A single scale factor is seen in equation (2), which does not produce information about the scale changes along each axis. One way to resolve this problem and deliver this information is to suppose that each axes have a various scale factor. By adopting this supposition, the above model becomes:

$$\mathbf{X}_2 = \Delta \mathbf{x} + \mathbf{R} \cdot \mathbf{S} \cdot \mathbf{X}_1$$

Where: S refers to the total scale matrix of the three scale factors:

(5)

$$\underset{3\times 3}{\mathbf{S}} = \mathbf{S}(\mu_1, \mu_2, \mu_3) = \begin{bmatrix} \mu_1 & & \\ & \mu_2 & \\ & & \mu_3 \end{bmatrix}, \quad (6)$$

$\mu 1$,$\mu 2$,$\mu 3$ are the three various factors of the scale for individual axis x, y, and z, those can be written as sum of unity and a scale change $\delta\mu j$, which often explicit as part of a million (ppm):

$$\mu_j = 1 + \delta\mu_j, \quad j = 1, 2, 3 \quad (7)$$

In this case the general equation (5) is not properly written, and some residues will evidence:

$$\mathbf{X}_2 - \varepsilon = \Delta \mathbf{x} + \mathbf{R}(\alpha_1, \alpha_2, \alpha_3) \cdot \mathbf{S}(\mu_1, \mu_2, \mu_3) \cdot \mathbf{X}_1 \quad (8)$$

Where: S refers to the residual vector / errors in the coordinates vector (X2):

$$\varepsilon = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{bmatrix}. \quad (9)$$

Previous Work

The production of reality-based 3D models is the representation of existent structures in various applications. However, for the complexity of urban inherent sites, multi-sensors approaches are desirable for 3D reconstruction and visualization of these sites [12].

In 2010, Amat *et al*. [13] produce 3D city modeling by combination aerial and CRP images. However, the method encountered a problem of assigning small buildings due to the scaling factor adopted. Therefore, they generated part of the model from CRP images only and later combined with the model created from aerial photos for visual applications and employments. Aerial images are still utilized to generate a 3D city model base, while CRP was used as a 3D reconstruction of the building facades. Further, objects like streets, trees and parking that can be put in a 3D city model can be generated from aerial images. The overall RMS amount of the building model was 0.005 meters and according to PhotoModeler standards, the RMS value must be less than 3 mm in order to obtain a good photographic project. On the other hand, [14] Checa et al., (2014) utilized the instance of the San Miguel Church in Agrida, Spain to create a methodology for data gathering and integration from different sensors. The case study was the architectural graphic documentation of an ecclesiastical building of the town of Ágreda (Soria) composed from several volumes. These where combined the acquisition of data through terrestrial and aerial photogrammetry with low-cost process. The "Phantom" quadrotor DJI with the "Gopro Hero 3" Silver edition camera was used to capture the aerial coverage. A total of 500 photographs have been used, creating their respective masks on the building to leave out cars and containers that were attached to one of the laterals of the church. The density of the generated point clouds was obtained from the facades oscillates which was 1,000,000 points, and 300,000 points from the air coverage. This density considered to be low but excellent for this type of work because it covers enough information to generate a good number of point clouds with optimal performance of the computer. Further Al-Adamat *et al*. [15], present their perspective in a photogrammetric workflow based on SfM technique to obtain 3D measurements of Ajloun Castle, Jordan from aerial and terrestrial photos. The 3D modeling framework consists of combining multi resolution photos from both overlapped low altitude and multi view terrestrial images. The acquisition of CRP images involves high resolution photos captured by a portable camera (Canon EOS 1100D), with a resolution of (4272 x 2848) pixels and (18 mm) focal length. The CRP photos utilized to document the walls of the castle with high precision and to generate wall structures with high details by matching dense images and aerial images of 18 overlapping aerial images covering the Ajloun Castle obtained from the archive of aerial photographic for Archaeology in the Middle East (APAAME) project.

Study Area Description

The study site was the University of Baghdad (UOB) campus with two case studies selected: the tower of the rectorship of the university; and the building of the Surveying Engineering Department (SED) within the college of Engineering campus, as shown in figure (2). The UOB is the largest university in Iraq, located in the center of Baghdad in Al-Jaderia area within Al-Karada district. It was built by the Iraqi government in the late 1950s and situated to the east bank of the Tigris River. The German Walter Gropius architect has designed the tower building architecture with height reaches to 75m above ground; however, the master plan began to build the faculties of engineering, science and contemporary arts for about 6,800 students in the 1960s. The University was extended in1982 to include 20,000 students with the supplement of other facilities. Later, the Iraqi Hisham N. Ashkurry architect and his colleague Robert Owen redesign the campus program. As the campus is considered one of the most important spots in Baghdad city with its iconic buildings, it was therefore selected to be virtually modelled in this study. This particularly important to keep it safe for next generations and also for any future architecture development as one of the most important academic cities in middle east.

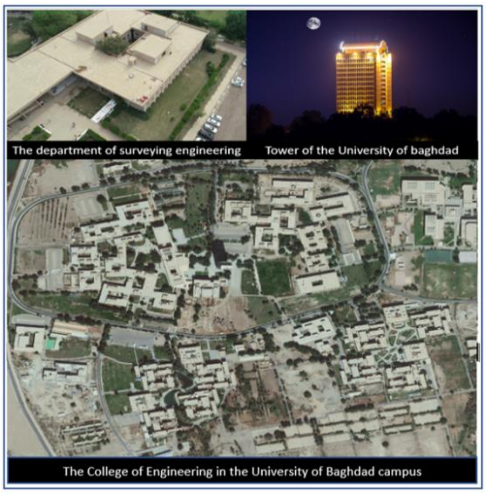

Fig-2: The University of Baghdad campus (UOB)

Methodological Framework

workflow was proposed to check performance and robustness of approaches and compare derivable towards achieving the objectives of this research. This was achievable using different algorithms from commercial software (Agisoft Photoscan), and comparsion between these results was made with validate results of ground truth data from traditional surveying. Field work involved planning, image capture, and control acquisition; however, office work involved photogrammetric data processing, integrating data, 3D point clouds extraction, and texture model generation. Planning process included the establishment of ground control points (GCP's) in study areas utilizing differential global positioning system (DGPS), target observation and 3D coordinates (X, Y, Z) derivation using Total station instrument, in addition to image capture. The terrestrial images were acquired at specific positions using static digital camera and also moving camera on a UAV platform. Through the processing stage, the 3D coordinates were computing automatically using multiple software's. This software were used to implement data integration process, triangulation, 3D point cloud generation, and 3D modeling and texturing. Figure (3) illustrate a diagram of the methodology proposed and applied in this research.

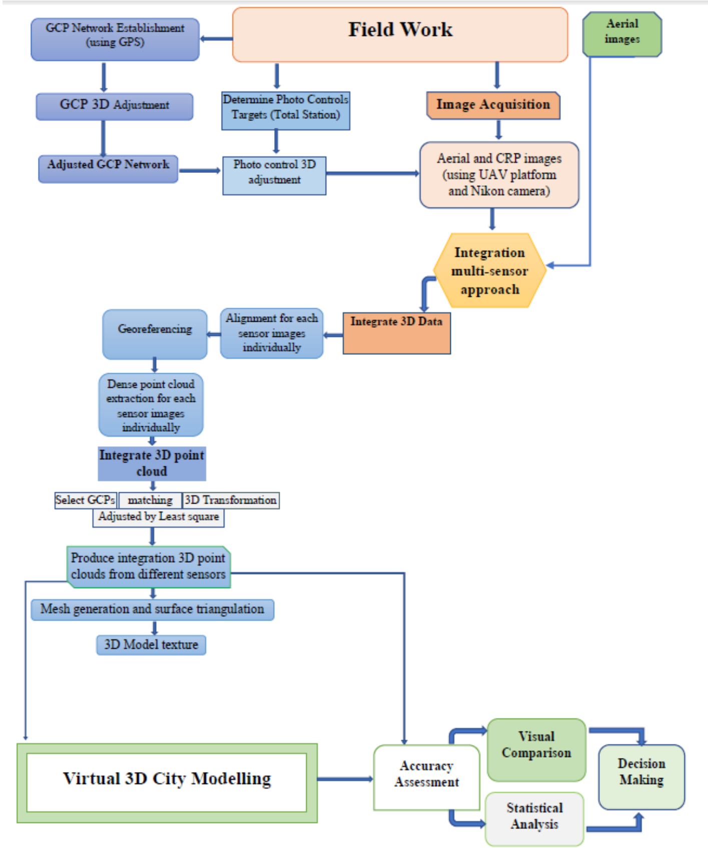

Fig-3: Workflow diagram of the research methodology

Data Collection and Preparation

To enhance the accuracy of the generated 3D models and verify the results delivered later, Ground Control Points (GCPs) network was created utilizing DGPS (Topcon GR5) in the study area, as shown in Figure (4). GCP points adjustment was applied using least squares approach in Topcon tools software. To connect the control points with any photogrammetry project, photo control targets must be designed and distributed above the surface of the object before capturing the images. In certain situations, it is difficult to put targets on the objects that lead to select photo-targets to look like distinctive features appear in common images, this is the case in UOB tower. However, in the case of SED building, the targets are designed and produced to be 10*10 cm. The targets measurements are implement using laser total station.

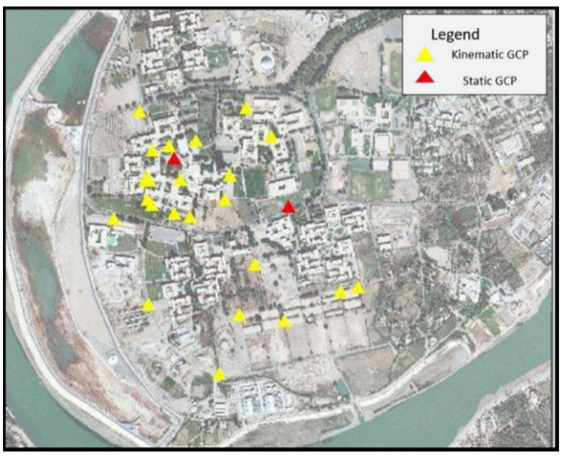

Fig-4: Distribution of the selected GCPs in UOB campus

In this study, two different camera sensors are used in the fieldwork to achieve the objective of the study. The first one is static digital camera (Nikon D610) with specifications shown in figure (5). The second camera used is DJI phantom 4 pro built-in camera within the unnamed aerial vehicle (UAV) platform. The specifications of DJI phantom 4 pro are shown in figure (6).

ISO Range	100 - 6400 in 1, 1/2 or 1/3 EV steps (50 - 25600 with boost)
Item Dimensions	5.55 x 3.23 x 4.45 in
Item Weight	1.87 lbs
Megapixels	24.3 megapixels
Optical Sensor Resolution	24.3 megapixels
Optical Zoom	3.5x
Photo Sensor Size	Full Frame (35mm)
Style Name	Body Only
Video Capture Resolution	FHD 1080p

Fig-5: The specifications of Nikon D610 camera

PHANTOM 4 PRO CAMERA	
Sensor	Effective pixels: 20M
Lens	FOV 84° 8.8 mm/24 mm (35 mm format equivalent) f/2.8 - f/11 auto focus at 1 m - ∞
ISO Range	100 - 3200 (Auto) 100- 12800 (Manual)
Photo	JPEG, DNG (RAW), JPEG + DNG
Mechanical Shutter Speed	8 - 1/2000 s
Electronic Shutter Speed	8 - 1/8000 s

Fig-6: The specifications of DJI Phantom 4 Pro camera.

For the Tower images, eight aerial images were utilized, two flight lines, each flight line contains four images and the block of the photos had 20% side lap and 60% end lap that cover an area about 1.333km^2. The flying height was 457m and the focal length was152.16mm. These aerial images are available as an aerial coverage to the case study site in digital form with 300 dpi resolution. However, the other aerial dataset consists of200 UAV images captured in 2018 (by DJI Phantom 4 Pro camera) for part of the university campus. The (B/H) ratio equals to approximate 1 with 9mm focal length and the distance between images in one vertical strip was 6m with 90 % overlap. The number of targets points was 146 (93 control points, 53 check points).

On the other hand, with SED building, the images used were 67 image from hand-held camera integrated with 86 UAV images. In close-range hand-held camera images, B/H ratio was 0.4 (B equal to 4.5 meters and H equal to 11.5 meters) and the focal length was 18 mm. In the UAV images, the flight plan consists of two flight lines with a third cross flight line. The images acquired at 60 meters high with 20° tilt from Nader (70° from the horizon). The overlap between images was 90 % and the side lap was 80% with focal length equals to 9mm. The number of total targets points was 66 (38 control points, 28 check points).

RESULTS AND ANALYSIS

3D Data Integration

The integration includes the combination of 3D dense point clouds generated from different sensors using Agisoft photoscane software. The process of integration includes creating two chunks in workspace.

Every chunk includes images from one sensor and performs alignment and dense point cloud process of individual chunks. Later, each chunk is tagged with the same tie-pass point targets then merge these two dense point cloud chunks with the help of tie-pass point targets to produce merged chunks. This contains a complete dense point cloud of the entire structure followed by meshing and texturing process. This integration performs on three case studies (UOB Tower, SED building in UOB campus and the college of Engineering (COE).

Dense Point Clouds and Texture Results

The dense point cloud process increases the intensity of the sparse point clouds those calculated during alignment processing. In this step, the object geometry was detected in more details as shown in figures (7), (8) and (9). The dense generation depended fundamentally on space intersection and technique of image matching. The exterior orientation parameters for each photo were utilized to estimate the 3D point cloud position. These parameters also performed a main role in matching digital photos. These dense point clouds could be adjusted and classified before exporting or reconstruction the 3D mesh model. In texture process, the texture from the original photos was re-projected to the surfaces of the generated model. If there was some specific surfaces that were shot from more than one photo, the projected texture of these photos would be mixed. The textures produce a high visual appearance and this requisite for virtual environments and other investigational applications, see figure (10).

Fig-7: The Dense point cloud of UOB Tower

Fig-8: The Dense point cloud of The SED building

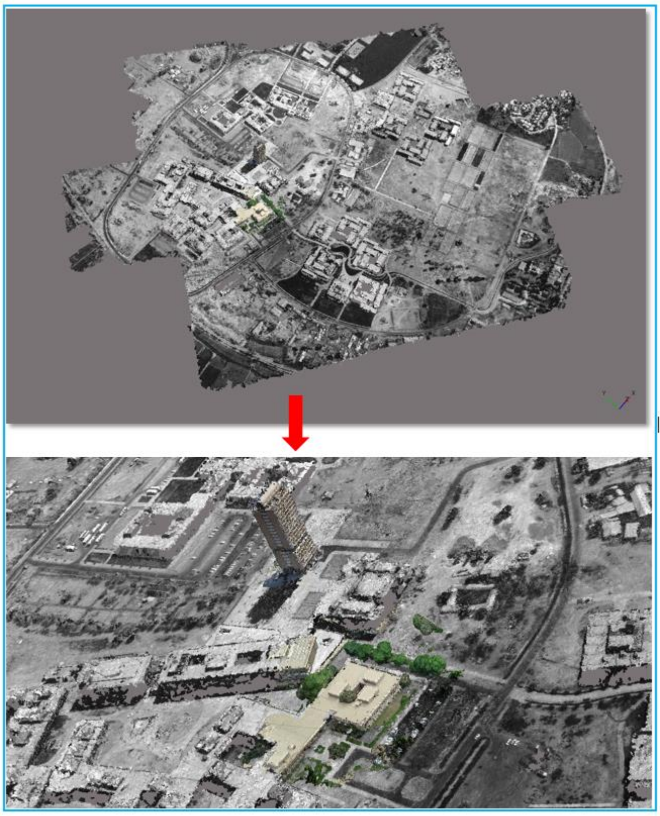

Fig-9: The Dense point cloud of the virtual city model of COE in UOB campus

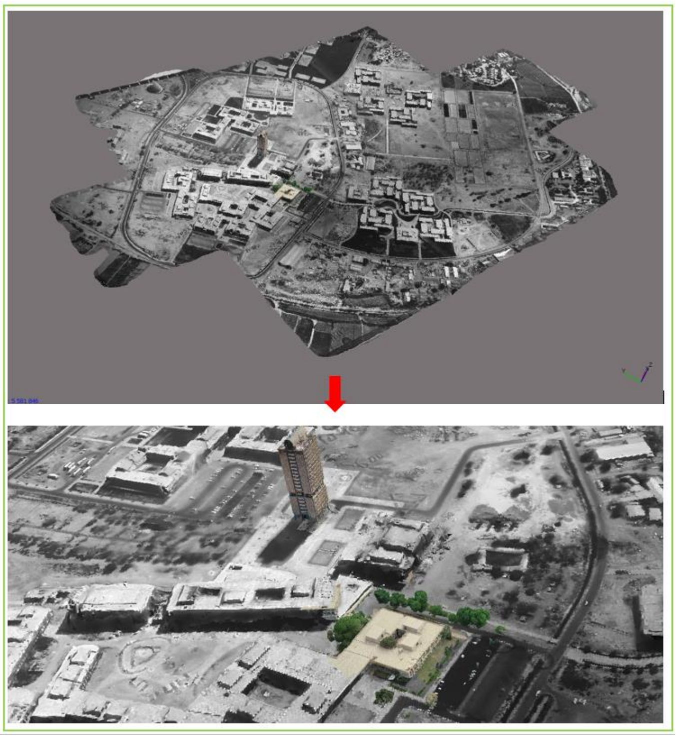

Fig-10: Texturing the virtual city model of the COE in UOB campus

Accuracy Validation

To evaluate the precision of the triangulation process in Agisoft Photoscan, RMSE was computed between the reference points obtained from adjusted GCPs and the computed points from the image triangulation process in all generated 3D models. The check points were marked in the models to provide conclusions. The RMSE values are produced in X, Y and Z directions with a number of check points are shown in the following Table.

Table-1: Summary of check point analysis for the recovered target points.

	No.of check points	*RMS (X) m*	*RMS (Y) m*	*RMS (Z) m*
	53	0.014081	0.015541	0.017733
SED building	28	0.010566	0.010708	0.013682
COE campus	11	0.271051	0.250703	0.301240

The result obtained from the 3D city model of UOB uses integration of multi-sensor data (aerial and close range images) with precise geometrical position and precise details with LOD3 for the UOB tower and the SED building.

CONCLUSIONS

This research identified the accomplishment and experimentation of a methodology presented to generate virtual city model to a certain study area from aerial and terrestrial images. For 3D modeling applications, integrating data from different image-based sources is effective and recommended, if data from single source is not available or even not enough to reach the required objective. SfM is a very good approach for professional processing of photogrammetric data, UAV photos processed by SfM approach can deliver accurate 3D models with high level of details (facades, floors, small geometry details) in contrary with static close-range imagery, if adequate capturing plan was adopted. The CRP method is characterized by the ability to solve aerial photos limitations in the development of 3D models of small structures. It also can make matching effective stereo photos depending on multiple overlapping photos and how can supply 3D models with adequate resolution approximately close to the sub-pixel level. The tilted aerial images proved suitable for facade texturing following image ortho-rectification process.

REFERENCES

1. Abdullah CC, Baharuddin NZ, Ariff MF, Majid Z, Lau CL, Yusoff AR, Idris KM, Aspuri A. Integration of Point Clouds Dataset from Different Sensors. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. 2017;42:9.
2. Zhang J. Multi-source remote sensing data fusion: status and trends. International Journal of Image and Data Fusion. 2010 Mar 1;1(1):5-24.
3. Gruen A, Huang X, Qin R, Du T, Fang W, Boavida J, Oliveira A. Joint processing of UAV imagery and terrestrial mobile mapping system data for very high resolution city modeling. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2013:175-82.
4. SMITH MJ, PRIESTNALL G, Asal F. Combining LIDAR and photogrammetry for urban and rural landscape studies. International Archives of Photogrammetry and Remote Sensing. 2000;33(B3):44-50.
5. Checa ZJ, Morales AF, Hernández LA. Combination of low cost terrestrial and aerial photogrammetry: three-dimensional survey of the church of San Miguel in Ágreda (Soria). Virtual Archaeology Review. 2014 May 2;5(10):51-8.
6. Blyenburgh PV. UAV's-current situation and considerations for the way forward. Development and Operation of UAVs for Military and Civil Applications. 1999.
7. Matthews NA. Aerial and close-range photogrammetric technology: providing resource documentation, interpretation, and preservation. US Department of the Interior, Bureau of Land Management, National Operations Center; 2008.
8. Brown LG. A survey of image registration techniques. ACM computing surveys (CSUR). 1992 Dec 1;24(4):325-76.
9. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE transactions on Medical Imaging. 1997 Apr;16(2):187-98.
10. Andrei O. 3D affine coordinate transformations. 2006.
11. Hofmann-Wellenhof B, Moritz H. Physical geodesy. Springer Science & Business Media; 2006 Oct 10.
12. Torres-Martínez JA, Seddaiu M, Rodríguez-Gonzálvez P, Hernández-López D, González-Aguilera D. A multi-data source and multi-sensor approach for the 3D reconstruction and visualization of a complex archaelogical site: The case study of Tolmo de Minateda. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. 2015;40(5):37.
13. Nor'Ainah Amat HS, Majid Z. Integration of aerial and close-range photogrammetric methods for 3d city modeling generation.
14. Checa ZJ, Morales AF, Hernández LA. Combination of low cost terrestrial and aerial photogrammetry: three-dimensional survey of the church of San Miguel in Ágreda (Soria). Virtual Archaeology Review. 2014 May 2;5(10):51-8.
15. Al-Adamat R, Al-Shawabkeh Y, Al-Kour O, Al-Shabeeb AR. A multi-resolution photogrammetric framework for digital geometric recording of large archeological sites: Ajloun Castle-Jordan. International Journal of Geosciences. 2016 Mar 2;7(03):425.