Al-Qadisiyah JournalFor Engineering Sciences 18 (2025) 265 – 271

Contents lists available at: http://qu.edu.iq

Al-Qadisiyah Journal for Engineering Sciences

Journal homepage: https://qjes.qu.edu.iq

Research Paper

A Hierarchical approach for efficient DTM and building footprint extraction from UAV images

Yousif A. Mousa [1]✉, ***Mourtadha Sarhan Sachit*** [2], ***and Ali Fadhil Hasan*** [3]

[1]*Civil Engineering Department, College of Engineering, Al-Muthanna University, Samawah, 66001, Iraq.*
[2]*Civil Engineering Department, College of Engineering, University of Thi-Qar, Nasiriyah 64001, Iraq.*
[3]*Department of Geography, College of Education for Humanity Sciences, Al-Muthanna University, Samawah 66001, Iraq.*

ARTICLE INFO

Article history:
Received 23 June 2024
Received in revised form 09 September 2024
Accepted 05 July 2025

keyword:
Drones
2D Regularization
Building Delineation
DSM
Digital Mapping

ABSTRACT

The utilisation of UAV imagery for the creation of digital maps is a compelling subject within the domains of photogrammetry and remote sensing. This work introduces a hierarchical method for automating the process of building, extracting, and outlining using images captured by drones. The flight plan should be initially planned to provide about 60-70% overlap to guarantee thorough coverage and precise image matching. The altitude of the drone should be adjusted based on the intended resolution to achieve a balance between capturing fine details and covering a larger region. Next, the technique of photogrammetric image matching was utilised to generate orthophotos and the Digital Surface Model (DSM). Moreover, the Digital Terrain Model (DTM) was extracted from the DSM to differentiate non-ground objects, including buildings. Subsequently, building segments were identified by applying a threshold to the difference between the Digital Surface Model (DSM) and the Digital Terrain Model (DTM), enabling accurate extraction of building segments. Finally, building polygons were generated involving two stages: coarse and refined, considering the least squares adjustment process to guarantee accuracy and detail. The proposed method was applied to drone images captured on the campus of Al-Muthanna University in the southwest of Iraq. The qualitative and quantitative investigation indicated that the building polygons obtained were highly promising, with approximately one-meter geometric accuracy. Nevertheless, accurately differentiating between buildings and other human-made structures (such as tents) and resolving issues related to mismatching error still pose significant difficulties, highlighting the need for additional investigation and development.

1. Introduction

Unmanned Aerial Vehicle (UAV) technology has garnered significant interest in several remote sensing applications, including urban planning, smart cities, environmental monitoring, and disaster management [1]. The main advantages are noticeable via offering high-resolution optical images at a comparatively affordable cost. This capability is further improved by the presence of sophisticated photogrammetric image matching software, which promotes the trend by producing orthophotos and digital surface models (DSMs) in a completely automated manner. The high-resolution photos are essential for generating extremely detailed DSMs, DTMs, and extracting building outlines [2]. However, the complex and unpredictable of building shapes in urban environments pose significant barriers. These issues require continuous efforts and additional research to enhance the utilisation of UAV technology in these complex environments, guaranteeing precise and dependable data for crucial applications. Recently, the use of UAV imagery for automatic DTM and building extraction has attracted a rising interest. The major advantages of employing these techniques compared to the classical mapping methods are demonstrated by their cost-effective [3]. Furthermore, [4] explained the enhancement of OpenStreetMap building regularization by using contour information utilizing oblique drone imageries. Several studies have examined automatic building extraction from UAV-based images and DSMs, emphasizing the significance of precise and automatic building footprint extraction via DSM and DTM data [5–7]. The research conducted by [8] emphasized the benefits and constraints associated with utilizing DSM and DTM for the extraction of three-dimensional (3D) building models. Furthermore, [9] examined the application of tilt photogrammetry to produce DSM, DTM, and normalized DSM (nDSM) to determine the building height using UAV point cloud data. With the availability of UAV optical images, the procedure of building detection and delineation starts with generating orthophotos and the Digital Surface Model (DSM). Alternatively, optical images can also produce point cloud data, which can then be segmented and used for building extraction, as described in [10]. Methods based on raster Digital Surface Models (DSMs) are particularly appealing due to the widespread availability of advanced image processing toolboxes [11]. The DSMs raster describes all visible man-made objects above ground, including buildings, trees, bridges, bridges etc. In contrast, Digital Terrain Models (DTMs) represent only the bare ground. Thus, applying a filter to differentiate between the DSM and DTM is essential for accurate building detection. Consequently, the so-called Normalized DSM is created by subtracting the DTM from the DSM, which helps in separating buildings in complex environments [12]. However, distinguishing between buildings and trees is challenging without the Normalized Difference Vegetation Index (NDVI), which typically differentiates vegetation from man-made objects. This limitation highlights the difficulty in accurately highlighting structures that rise above the ground level.

*Corresponding Author.
E-mail address: yousif.mousa@mu.edu.iq ; Tel: (+964) 783-888 7954 (Yousif Mousa)

https://doi.org/10.30772/qjes.2024.151126.1284

Nomenclature

AI	Artificial intelligent	*DTM*	Digital terrain model
3D	Three-dimensional	*GIS*	Geographic information system
2D	Two-dimensional	*NDVI*	Normalized difference vegetation index
CNN	Convolutional neural network	*NGP*	Network ground points
DSM	Digital surface model	*UAV*	Unmanned aerial vehicle

Furthermore, objects such as buildings and trees that have significant height above the ground are identified by applying a height threshold, commonly set at around 2.5 meters [13]. Some research has considered using a planarity map to remove trees by analysing the height variation within the DSM raster [14]. The planarity concept is based on the fact that buildings tend to have more uniform, planar surfaces, compared to the irregular surfaces of trees. By analyzing these height variations, trees can be effectively eliminated, resulting in a binary image that contains only building segments. Consequently, the final output of this process should be a raster image exclusively depicting building segments. For practical applications, the raster representation of building segments needs to be transformed into a vector format. This transformation process is known as vectorization [15]. During vectorization, the detected building segments are converted into vector data, typically in the form of polygons. These polygons represent the outlines of the buildings and are stored in a shapefile format. Shapefiles are widely used in geographical software, such as Geographic Information Systems (GIS), for various spatial analyses and mapping purposes. This process of converting building segments into vector format and creating shapefiles is also referred to as building regularization [16]. Building regularization ensures that the building outlines are accurately represented and can be easily integrated into GIS applications. This enables further analysis, such as analysing spatial relationships, and integrating with other geospatial datasets for urban planning, disaster management, and other applications. In the context of building regularisation, recent trends show that state-of-the-art machine learning models are being used to generate Digital Terrain Models (DTMs) from UAV datasets [17] which can be used for building detection and regularisation. Alongside, the Convolutional Neural Network (CNN) based methods have made significant developments. These methods have been integrated as plugins into commonly utilized platforms like QGIS, with notable examples including mapflow [18] and Polymapper, introduced by [19]. However, challenges remain in generating precise edges and handling occlusions due to reliance on curve initialization [20].

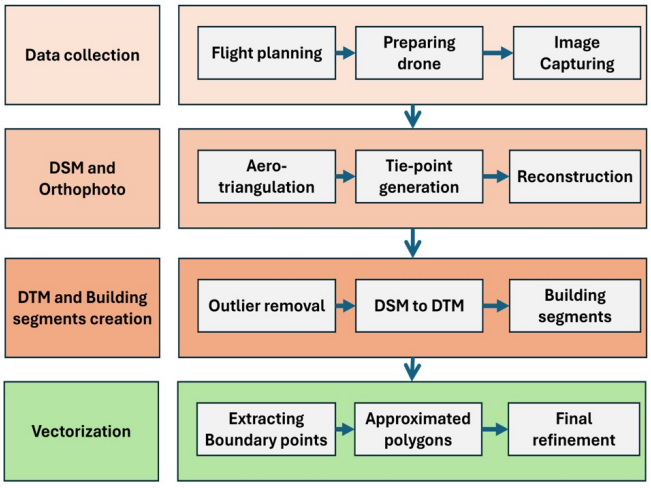


Figure 1. The proposed Methodology. The first stage was data collection, including flight planning, preparing the drone, and image capturing. Next, generating orthophotography and DSM. Further, extracting DTM and Building segments. Finally, the raster to vector transformation stage (vectorization).

While the Artificial intelligent (AI) based models (e.g., mapflow) can be effective, they necessitate sufficient training data and a cognitive approach to the process of extracting valuable information [21]. Additionally, existing building outlining methods still face difficulties in handling complex building shapes with holes [1]. Overall, drone technology has proven significant in efficiently and economically acquiring high-resolution datasets for building footprint extraction. Consequently, there has been a rise in the demand for utilizing UAV imagery and DTM data for building extraction and 3D modelling, showcasing the potential benefits and challenges associated with these technologies. In this research, the use of UAV images for the purposes of building detection and regularization is investigated. Furthermore, the study explains key stages, including Digital Surface Model (DSM) generation, Digital Terrain Model (DTM) extraction, building detection, and 2D reconstruction. Finally, a qualitative and quantitative analysis are conducted to evaluate the proposed method and to highlight the challenging scenarios. This paper is organized as follows. Methodology is explained in the next section including data collection, orthophoto and DSM generation, DTM and building extraction, and outlining. In Section 3, the obtained results will be presented and analysed. Lastly, the research conclusion will be given in section 4.

Figure 2. Flight plan describing front and side overlap, camera angle, and mapping out the drone's path.

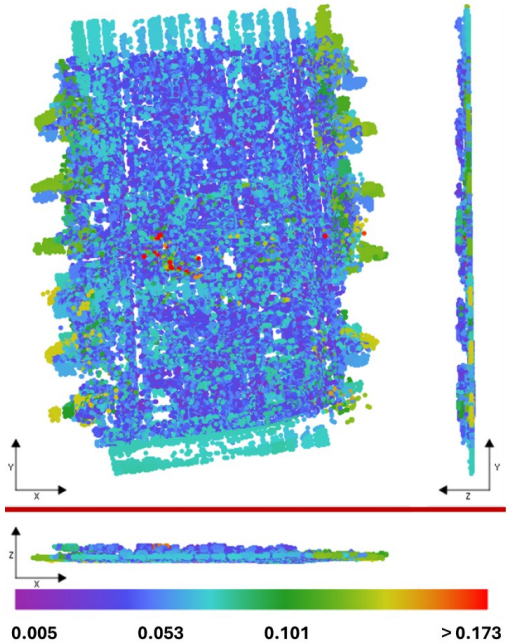
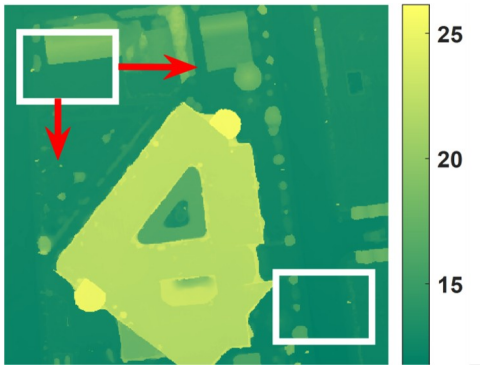


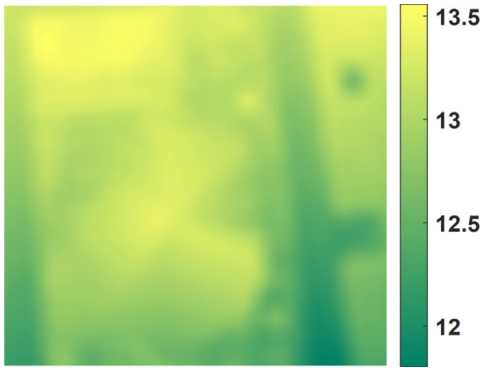
Figure 3. . Positional accuracy: Displays of all tie points are shown from the top view (XY plane), side view (ZY plane), and front view (XZ plane), with colours indicating the uncertainty of each point's position. The uncertainties are measured in meters, ranging from a minimum of 0.014 *m* to a maximum of 0.5855 *m* , with a median position uncertainty of 0.0531 *m*.

2. Methodology

Figure 1 below illustrates the process of creating building polygons from drone images through a hierarchical series of stages. It starts with the data collection stage, including flight planning to map out the drone's path and camera settings, and image acquisition.



(a) DSM



(b) DTM

(c) Building segment

Figure 4. DTM and building segments extraction process. a) DSM raster, b) DTM of building, and c)the extracted building segments. The white box (in a) indicates the window for DTM extraction filter. It moves across the whole image raster.

The process then moves to creating the DSM and Orthophoto from the captured overlapped image, including aerial triangulation to calculate camera positions, tie point creation to identify common points in overlapping images, and reconstruction to create the Digital Surface Model (DSM) and orthophotography images. The context capture software is used for creating DSM and orthophotography images. Next, the stage of DTM and the creation of building segments, which involves removing outliers to discard erroneous values in the DSM raster, extracting of digital terrain model (DTM) from the DSM, and identifying building segments. Finally, the routing phase begins by defining the boundary points of the building segments, followed by creating approximated polygons, and refinement of these polygons for accuracy and detail. The proposed methodology was applied to investigate its effectiveness on drone images.

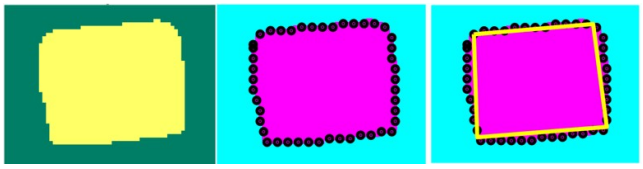


Figure 5. Building polygons creation. The building segment is on the left side, the initial boundary (black dots) are presented in the middle, and the created building polygon is on the right side (yellow).

2.1 Study area and data Collection

The chosen study area (0.15 km^2) is located on the campus of Al-Muthanna University South-West of Iraq within 527300E to 527600E, and 3466900N to 3467350N. It consists of several buildings with very complex structures and shapes. Low bushes and grass are the main green area in the campus. Firstly, the flight plan was prepared as shown in Fig. 2. Front and side overlap were set to 70% with flying height approximately 150m. The flight condition for the drone would be in the middle of the day to reduce the shadow effects and to improve images quality. It is preferable for the temperature to be moderate, e.g., less than 40°. Light winds is an important factor for a stable flight to reduce errors in calculating the camera orientation parameters. The interior camera calibration parameters are provided in Table 1. These parameters are, focal length, the principal points positions (X, Y), lens distortion coefficients (K1, K2, K3) and tangential distortion. These parameters were estimated based on the concept of the collinearity equations. The optimization process changed these parameters, indicating an improvement in camera calibration accuracy by adjusting the focal length and distortion parameters. Images were captured using Drone type "Mavic 2 Pro". The GNSS system, incorporating GPS and GLONASS, delivers the following hovering accuracy: vertically, $\pm 0.1m$ when vision positioning is enabled and $\pm 0.5m$ with GPS positioning; horizontally, $\pm 0.3m$ with vision positioning enabled and $\pm 1.5m$ with GPS positioning, see http://www.dji.com/mavic-2 for more detail. A total of 346 images were successfully calibrated. The number of created tie points was 122822 points, with a median of 1463 points per photo. The Tie points were detected automatically using the Context capture software. The median position uncertainty equals 0.0531 meters with a minimum uncertainty of 0.014 m and a maximum of 0.5855 m as shown in Fig. 3.

Table 1. Camera calibration parameters.

Value	Focal length	Principal point X	Principal point Y	$K1$	$K2$	$K3$
Previous	10.26	2665.25	1483.84	-0.006	-0.005	0.040
Optimized	13.12					

2.2 Orthophoto and DSM generation

For the creation of orthophoto images and DSM, the Context Capture software (Academic version) was utilized https://www.bentley.com/software/. The software contains many tools and is able to process images and produce orthophotography, DSM, point cloud data, and much more. Additionally, the software is capable of doing some other calculations, such as volume. It is able to manage small holes in the DSM by interpolation. The resolution of the created orthophotography and DSM was 0.05m. This very high resolution allows a great chance for visible inspection as well as providing a very detailed DSM.

2.3 DTM and Building segments extraction

The DSM is a raster image describing all visible man-made objects on the ground surface, including buildings and trees as well as the bare ground earth. The pixel values refer to the elevation. While the DTM represents only the bare ground part. In order to identify building segments, DTM must be separated first. In this paper, the so-called Network Ground Points (NGP) method is considered Error! Reference source not found. Firstly, outliers (extremely low or high pixel values) should be eliminated. The NGP algorithm aims to detect seed ground points by applying a window filter containing multi-directional scanlines. The scanline is defined as a vector of data extracted from the DSM

image within the dimensions of the applied filter. The filter moves across the entire DSM image at a predetermined shift in both X and Y directions. Figure 4 illustrates DSM raster (top), DTM in the middle, and the extracted building segments in the bottom. The filter window is indicated by the white box on the DSM (Fig. 4). The filter moves across the whole image horizontally and vertically, looking for low-elevation pixels as seed ground points. The recognized seed ground points are then utilized to generate the DTM through an interpolation technique. Further details were given in [22]. Furthermore, the so-called Normalized Digital Surface Model (nDSM) is created by subtracting DTM from DSM. This nDSM contains elevated man-made objects above ground, including buildings and trees, for example. Finally, building segments were identified by thresholding the resulting nDSM. For instance, only objects have significant height above ground (e.g., 3 meters) were detected as buildings. Separated trees (e.g., low bushes) were eliminated using structural component analysis based on their size.

2.4 Building polygons creation

After generating building segments, the vectorization stage is applied to convert the result from raster format into vector format. This is important because geographical applications such as Geographic Information System (GIS) deal with vector data more professionally and practically in terms of modifying, updating, and adding attributes. The vectorization part involved two stages: generating approximated building polygons, followed by refinement. Firstly, the object component algorithm is used to extract border points located at the outline of building segments. The algorithm begins by identifying the outer edges of the objects and the edges of any holes within those objects. The image used must be binary, with non-zero pixels representing the objects and zero pixels representing the background. The extracted boundary is represented by coordinates, which are organized in a clockwise or anti-clockwise direction. Figure 5 describes the process of building polygons creation. The building segment is on the left side, the initial boundary points (black dots) are presented in the middle, and the created building polygon is on the right side (yellow). so that a list of boundary points that are indexed by their row and column as coordinates is collected. Then, this list of boundary points is reduced by keeping only corner points as vertices and eliminating all redundant points located between the selected vertices using the likelihood function introduced in [23] as given in Eq. 1.

$$L_i = A_i + a \times D_i \times sin(\theta_i) - b \times E_i^2 \tag{1}$$

The given equation sequentially computes the values (L_i) for each boundary point starting from the initial point. Here, (A_i) denotes the area of simplified polygons formed from the starting point to the current point and back to the starting point. (D_i) represents the distance from the starting point to the current point. At each boundary point, the sine of (θ_i) is calculated. (E_i^2) is the mean square of the orthogonal distances from the boundary points between the starting point and the current point. The constants (a) and (b) are set experimentally to 20.0 and 2.0, respectively, to balance the weighting of these different terms. The proposed methodology, including DTM extraction and building outlining, was implemented using the programming platform (MATLAB). Further details can be found in [23]. The results represent approximated solutions for creating building polygons. Therefore, the final refinement step was applied to enhance the geometrical accuracy of the created polygons using least squares adjustment. The aim is to perform a best fit between the border points with respect to their assigned polygon sides. Additionally, orthogonality constraints are considered by generating right angles in rectilinear buildings.

3. Results and discussion

The proposed procedure was tested on part of Al-Muthanna University campus as shown in Fig. 6. The process of building regularization is fully automated, eliminating the need for any manual digitizing, which makes it look very promising. This degree of automation suggests that it can be highly efficient and reliable, potentially leading to significant improvements compared to traditional digitizing. Figure 6 shows the generated orthophoto image (created by the Context Capture software) overlaid by the final created building polygons (MATLAB implementation). The figure showing a digital map contains buildings, tents, and other man-made objects that are outlined. More qualitative and quantitative analyses focusing on challenging scenarios are provided in the next section.

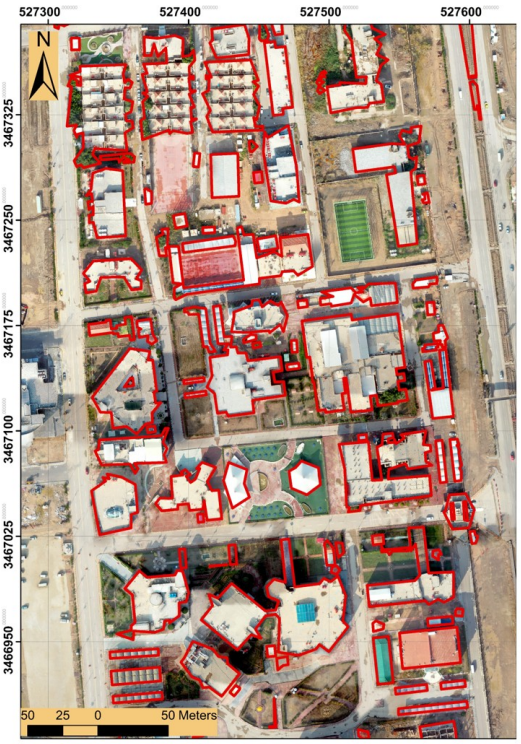
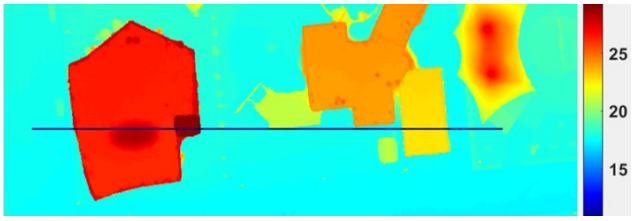
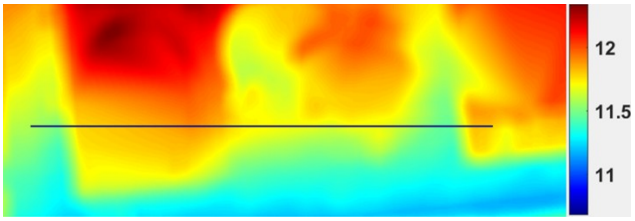


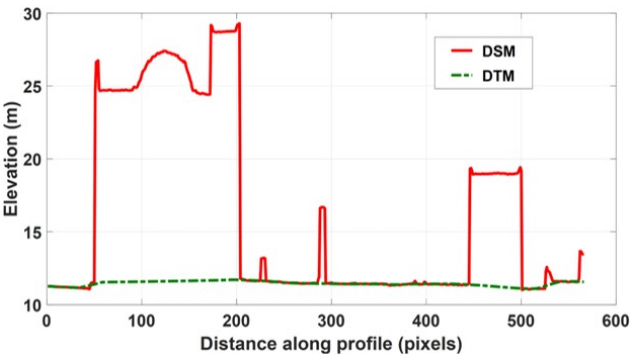
Figure 6. Final created digital map. It is showing an orthophotography image overlaid by the created building polygons (red polygons).



(a) DSM



(b) DTM



(c) Combines DSM and DTM

Figure 7. Section of the extracted DTM, where a) DSM part, b) DTM part, and c) A profile combines DSM and DTM.

3.1 Qualitative analysis

The results of the proposed framework are qualitatively analysed by highlighting some challenging spots. Firstly, the extracted DTM is investigated. Figure 7 provides a section from the extracted DTM (in the middle) from the input DSM (at the top). A profile that combines both DTM and DSM was plotted at the bottom of the figure. Pixel heights in the DSM were approximately between 10 to 30m. After eliminating elevated man-made objects (Buildings) and filling their spots by an interpolation technique, the DTM was created. The heights in the created DTM fluctuated in the range of approximately 11 to 12 meters. At the bottom of Figure 7, a detailed comparison between the extracted digital terrain model (DTM) and the actual land surface represented in the input digital surface model (DSM). The presented profile shows that the DTM marked by the green dashed line closely adheres to the actual land surface features recorded in the DSM. This close alignment indicates that the DTM extraction process was successful in accurately representing the underlying terrain and removing man-made objects such as buildings.

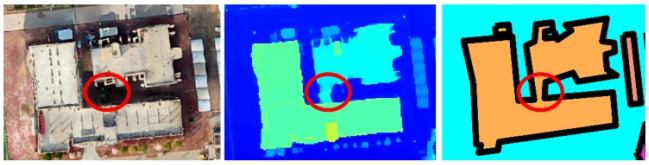


Figure 8. Difficult example presented two connected buildings due to a mismatching error caused by occlusion.

In the context of building footprint extraction, some challenging scenarios were highlighted. For instance, Fig. 8 provides orthophotography image (left), DSM (middle), and the created building polygon (right). It can be seen that the two buildings in the orthophotography are separate from each other. While the created DSM presented a narrow connection between them, as indicated by the red circle. This refers to the mismatching problem caused by occlusion and/or shadow effects. Therefore, these two buildings are regularized as one complex shape, as shown on the right. The other challenge facing the applied framework is in terms of distinguishing between buildings and other man-made objects like tents and sheds. This is because the algorithm is based on the height difference between DTM and DSM. Thus, all objects having more than 3m are classified as buildings. Figure 9 illustrates this issue as the two objects (tents) were identified and regularized as buildings. Figure 10 below presents a complex non-rectilinear building as shown in the top (orthophoto) and the final extracted polygon at the bottom (red line). Additionally, the building contains a hole or inner courtyard in the middle. The applied regularization method overcame such a complex case, even with inner holes. Figure 10 shows the boundary points (black dots) and the created building polygon (red). On one hand, it can be seen that the generated polygon follows the outlines of the building segment precisely. Furthermore, the best-fitting of the created polygon with respect to the boundary points is noticeable. This is the advantage of applying the least square adjustment, in which the shifting between extracted polygon edges and the input boundary points is minimized. On the other hand, certain fine details of the building were oversimplified in the created polygon. For example, the building includes complicated architectural features, such as a half-circle structure, highlighted by the red arrow in Figure 10. Due to the complexity of these features, the created polygon was unable to accurately capture and represent them, resulting in a loss of detail. This oversimplification occurred because the process of converting the building's geometry into a polygon format was unable to accommodate such complex shapes, leading to a more generalized and less precise representation.

3.2 Quantitative analysis

For evaluation purposes, a quantitative analysis was performed. To achieve this task, a ground truth dataset was created by manually digitizing the outlines of the existing buildings in the scene. As a measure, the Root Mean Squared Error (RMSE) is used as a standard topographic metric for evaluation. Each extracted building polygon is compared to its assigned reference polygon from the digitized ground truth data. It is known that the number of vertices in the extracted and reference models is not equivalent. Thus, the vertex-to-vertex evaluation model should be avoided. To achieve a reasonable assessment, the shift from each vertex in the extracted model to its closest vertex or edge in the reference model is calculated and vice versa. Figure 11. presents the created polygons (light green) and the reference polygons (red) overlaid with an orthophotography image (top) and with DSM (middle). The bottom row describes the displacements between the created and the reference polygons. These displacements were estimated as described above (vertex to closest vertex or edge). The calculated RMSE was approximately one meter which is slightly high. The main reason for that is related to some remained trees as well as some existing shelters in contact with buildings. The red arrows (Fig. 11) indicate a shelter that is attached to a building. While the reference polygon excluded this shelter as it was digitized manually, it is included in the created polygon because it is elevated over the ground and in contact with the building. Consequently, the calculated RMSE becomes higher. However, the extracted edges on the other sides followed the reference polygon precisely.

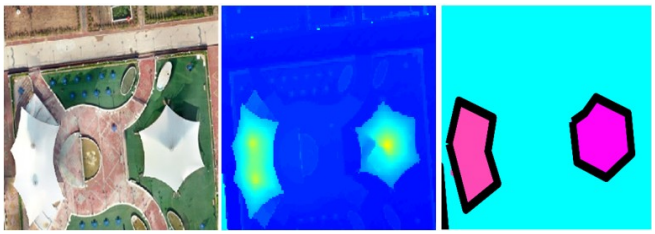
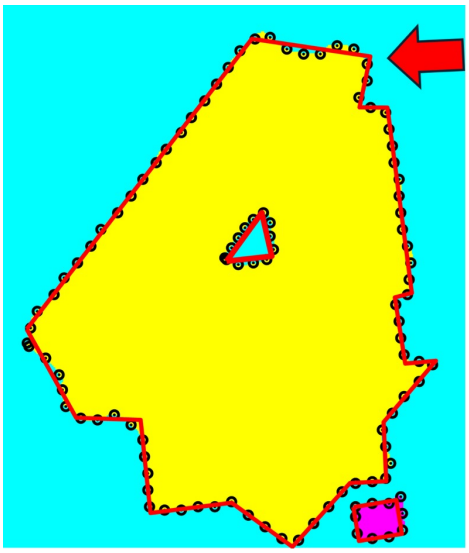


Figure 9. High tents classified incorrectly as buildings.

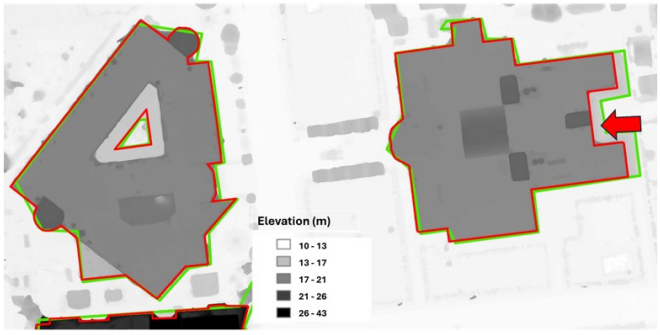
(a) Study building with a hole.



(b) Boundary points.

Figure 10. Complex building example contains a hole. Boundary points were indicated by small black dots, and the created building polygons are coloured red.

(a) Study building with a hole.



(b) Boundary points.

(c) Combines DSM and DTM

Figure 11**.** Positional accuracy assessment of the extracted polygons. The created polygons (green) and the reference polygons (red) overlaid with a) Orthophoto image, b) DSM, and c) Row describe the shift distances between the created and the reference polygons.

The extracted building polygons seem very promising,g which can be further used in several applications such as urban planning and 3D city modelling. For instance, these polygons could be exported as shapefiles and used in geographical software such as GIS for the purposes of updating existing map and change detection. The main challenge that faced the implemented methodology was the elimination of trees using the DSM only. The drone images involved three bands, Red-Green-Blue (RGB). For instance, the near-infrared (NIR) band was not available. Separated trees can be removed using a connected component analysis technique based on their area size. However, some of high high-elevation trees seem very hard to eliminate due to their large area, especially those in touch and/or overhanging with buildings. Therefore, the geometric accuracy of the extracted building polygons was reduced. With the absence of the NIR band, the DSM and RGB images could be fused for further enhancement.

4. Conclusion

The potential and practical use of UAV technology has been proven by providing high-resolution images at relatively low cost. Alongside, the photogrammetric image matching software supports this trend by creating orthophotography and DSM fully automatically. In this study, a detailed procedure for generating building polygons from raw RGB images captured by a drone was proposed and investigated. The obtained results were evaluated qualitatively and quantitatively to prove the capability of the applied methodology. As geometric accuracy, the created buildings' polygons achieved approximately one-meter RMSE. The significance of this study is that the manual work of creating building polygons from drone images is minimized in favour of a fully automatic procedure. While it succeeded in regularizing complex buildings, even those ones with courtyards, it was unable to differentiate between buildings and other man-made objects such as tents. Nevertheless, mismatching errors caused by occlusion and/or shadow, as well as overhanging trees, are the most challenging scenarios in the procedure of building polygon generation. A procedure that fused RGB images with DSM for tree removal could be a future research direction.

Authors' contribution

All authors contributed equally to the preparation of this article.

Declaration of competing interest

The authors declare no conflicts of interest.

Funding source

This study didn't receive any specific funds.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] Q. Li, L. Mou, Y. Sun, Y. Hua, Y. Shi, and X. Zhu, "A review of building extraction from remote sensing imagery: Geometrical structures and semantic attributes," *IEEE Trans. Geosci. Remote Sens*, vol. 62, p. 1–15, 2024. [Online]. Available: https://doi.org/10.1109/TGRS.2024.3369723

[2] W. Liu, M. Yang, M. Xie, Z. Guo, E. Li, L. Zhang, and D. . Wang, "Accurate building extraction from fused dsm and uav images using a chain fully convolutional neural network," *Remote sensing*, vol. 11, no. 24, p. 2912, 2019. [Online]. Available: https://doi.org/10.3390/rs11242912

[3] F. Remondino, L. Barazzetti, F. Nex, M. Scaioni, and D. Sarazzi, "Uav photogrammetry for mapping and 3d modeling: Current status and future perspectives," *In Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g): 14-16 September (2011), Zurich, Switzerland (pp. 25-31). International Society for Photogrammetry and Remote Sensing (ISPRS)*, vol. 100, no. 12, p. 3450–3469, 2022. [Online]. Available: https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-25-2011

[4] X. Zhuo, F. Fraundorfer, F. Kurz, and P. Reinartz, "Optimization of openstreetmap building footprints based on semantic information of oblique uav images," *Remote Sensing*, vol. 10, no. 4, p. 624, 2018. [Online]. Available: https://doi.org/10.3390/rs10040624

[5] Z. Farajzadeh, M. Saadatseresht, and F. Alidoost, "Automatic building extraction from uav-based images and dsms using deep learning. isprs ann. photogramm," *Remote Sens. Spat. Inf. Sci*, vol. 10, no. 1-3, p. 171–177, 2023. [Online]. Available: https://doi.org/10.5194/isprs-annals-X-4-W1-2022-171-2023

[6] J. Zhou, Y. Liu, G. Nie, H. Cheng, X. Yang, X. Chen, and L. Gross, "Building extraction and floor area estimation at the village level in rural china via a comprehensive method integrating uav photogrammetry and the novel edsanet," *Remote Sensing*, vol. 14, no. 20, p. 5175, 2022. [Online]. Available: https://doi.org/10.3390/rs14205175

[7] A. Shukla and K. Jain, "Automatic extraction of urban land information from unmanned aerial vehicle (uav) data," *Earth Science Informatics*, vol. 13, no. 4, pp. 1225–1236, 2020. [Online]. Available: https://doi.org/10.1007/s12145-020-00498-x

[8] E. Fissore and F. Pirotti, "Dsm and dtm for extracting 3d building models: advantages and limitations. international archives of the photogrammetry," *Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 1539–1544, 2019. [Online]. Available: https://doi.org/10.5194/isprs-archives-XLII-2-W13-1539-2019

[9] M. Sharma and R. Engineering, "Building footprint extraction from aerial photogrammetric point cloud data using its geometric features," *Journal Build. Engineering*, vol. 76, p. 107387, 2023. [Online]. Available: https://doi.org/10.1016/j.jobe.2023.107387

[10] Y. Dai, J. Gong, Y. Li, and Q. Feng, "Building segmentation and outline extraction from uav image-derived point clouds by a line growing algorithm," *International Journal of Digital Earth*, vol. 10, no. 11, pp. 1077–1097, 2017. [Online]. Available: https://doi.org/10.1080/17538947.2016.1269841

[11] Kokalj and R. Hesse, "Airborne laser scanning raster data visualization: a guide to good practice," *Založba ZRC*, vol. 14, no. 1, pp. 51–57, 2017. [Online]. Available: https://doi.org/10.3986/9789612549848

[12] X. Huang, W. Yuan, J. Li, and L. Zhang, "A new building extraction postprocessing framework for high-spatial-resolution remote-sensing imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 2, pp. 654–668., 2016. [Online]. Available: https://doi.org/10.1109/JSTARS.2016.2587324

[13] M. Awrangjeb, C. Zhang, and C. S. Fraser, "Building detection in complex scenes thorough effective separation of buildings from trees," *Photogrammetric Engineering Remote Sensing*, vol. 78, no. 7, pp. 729–745, 2012. [Online]. Available: https://www.isprs.org/resources/datasets/benchmarks/UrbanSemLab/results/papers/Awrangjeb_Building_2012.pdf

[14] K. F. West, B. N. Webb, J. R. Lersch, S. Pothier, J. M. Triscari, and A. E. Iverson, "Context driven automated target detection in 3d data," *Automatic Target Recognition XIV. SPIE*, vol. 5426, p. 133–143, 2004. [Online]. Available: https://doi.org/10.1117/12.542536

[15] M. Heitzler and L. Hurni, "Cartographic reconstruction of building footprints from historical maps: A study on the swiss siegfried map," *Transactions in GIS*, vol. 24, no. 2, pp. 442–461, December 2020. [Online]. Available: https://doi.org/10.1111/tgis.12610

[16] S. Wei, T. Zhang, S. Ji, M. Luo, and J. Gong, "Buildmapper: A fully learnable framework for vectorized building contour extraction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 197, pp. 87–104, 2023. [Online]. Available: https://doi.org/10.1016/j.isprsjprs.2023.01.015

[17] C. Gevaert, M. Persello, N. C., and V. F., "2018, u. a deep learning approach to dtm extraction from imagery using rule-based training labels," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 142, p. 106–123, 2018. [Online]. Available: https://doi.org/10.1016/j.isprsjprs.2018.06.001

[18] Mapflow, "Mapflow.ai - ai mapping and imagery analysis platform, version 2.5.0. mapflow.ai." 2024. [Online]. Available: https://mapflow.ru/

[19] Z. Li, J. D. Wegner, and A. Lucchi, "Topological map extraction from overhead images," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, p. 642–665, 2019. [Online]. Available: https://doi.org/10.48550/arXiv.1812.01497

[20] W. Zhao, C. Persello, and A. Stein, "Building outline delineation: From aerial images to polygons with an improved end-to-end learning framework," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, no. 2021, p. 119–131, May 2021. [Online]. Available: https://doi.org/10.1016/j.isprsjprs.2021.02.014

[21] W. Liao, X. Lu, Y. Fei, Y. Gu, and Y. Huang, "Generative ai design for building structures," *Automation in Construction*, vol. 157, p. 105187, 2024. [Online]. Available: https://doi.org/10.1016/j.autcon.2023.105187

[22] Y. A.-k. Mousa, P. Helmholz, and D. Belton, "New dtm extraction approach from airborne images derived dsm," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-1/W1, p. 75–82, May 2017. [Online]. Available: https://doi.org/10.5194/isprs-archives-XLII-1-W1-75-2017

[23] Y. A. Mousa, P. Helmholz, D. Belton, and D. Bulatov, "Building detection and regularisation using dsm and imagery information," *The Photogrammetric Record*, vol. 34, no. 165, pp. 85–107, 2019. [Online]. Available: https://doi.org/10.1111/phor.12275

How to cite this article:
Yousif A. Mousa, Mourtadha Sarhan Sachit, and Ali Fadhil Hasan. (2025). 'A Hierarchical approach for efficient DTM and building footprint extraction from UAV images', Al-Qadisiyah Journal for Engineering Sciences, 18(3), pp. 265-271. https://doi.org/10.30772/qjes.2024.151126.1284