ORIGINAL ARTICLE

A Grid-based Automated Building Extraction Technique for Low-cost UAV Images

Gamage, R. R.*, Nalani, H. A.

Department of Remote Sensing and GIS, Faculty of Geomatics, Sabaragamuwa University of Sri Lanka, Sri Lanka

ARTICLE INFO

Article history:
Received 12 April 2021
Received in revised form 22 May 2021
Accepted 23 May 2021
Available online 31 May 2021

Keywords:
Building extraction
Photogrammetry
Low-cost UAV
UAV images
Point cloud

ABSTRACT

With increasing urbanization, new technology is required to fulfil both human and environmental needs. At present, low-cost UAVs are used in surveying and mapping, and during the past few years, they have reached a level of practical requirements to allow the use of these systems as mapping platforms. Moreover, UAV based mapping provides required accuracy in line with cadastral laws and policies. Extraction of urban objects is a pre-requisite in various applications. In general, detection of buildings plays a major role in the field of remote sensing image processing, and also in urban planning and management. However, there is no 'proper' method developed to detect building features automatically from UAV images because there are usually too many details and distortions on the images. This paper presents an effective approach for extracting buildings from UAV images through the incorporation of orthophotographs and dense point clouds, rather than the traditional pixel based classification. In this method, different feature-based conditions are introduced with the help of a grid-based data structure for more accurate and quick extraction of building features. To verify the generality and advantage of the proposed method, the procedure is evaluated by performing experiments with a dataset acquired over the study area, which has a variety of building patterns and styles. The experimental results show an excellent performance in the detection of buildings, with an average overall accuracy greater than 80%. The final overall correctness and quality of building extraction are more than 80% and 65%, respectively. Therefore, there is a need to focus on more advanced conditions for building detection, to obtain optimum results.

1 Introduction

In photogrammetry, it is considered to be quite accurate if the consumer meets the exact requirements for preparing the 3D model. It is quite easy to use photogrammetry in those places where an object is to be measured, and to see how much dirt need to be cleared, particularly if it is all bare soil. When compared to the information coming from a human, the former is quite accurate. Budget constraints are, as ever, what they are and if the limits are too high, a project may not even come into being. That can force people, regardless of the other specifics, to choose the most economically viable alternative. Many people have to choose that, but, pursuing a low-cost approach will involve risks as to the nature of the specific production or deliverable. Participants must be aware of the risks, and it is important to properly assess how the technologies that you use can influence a project's possible success.

The most recent photogrammetric technique, i.e. Unmanned Aerial Vehicle (UAV), has more advantages due to the low time and cost consumption for both the smaller and large scale studies (Nex and Remondino, 2014; Ramon Soria et al., 2016). Further, the acquired images are mostly of high resolution, with the accuracy ranging from a sub meter level to centimeter level (Gerke and Przybilla, 2016; Harwin and Lucieer, 2012). Moreover, using a drone containing multiple GCPs, absolute precision in 5-10 cm can be easily achieved for smaller survey areas, which is good enough in most cases (Yao et al., 2019). Therefore, it has been a most prominent data acquisition platform which benefits a wide range of applications such as urban monitoring, land use analysis, building reconstruction, 3D city modeling, disaster and real estate management. These high resolution UAV images can be used to extract urban features effectively (He et al., 2019). Thus, the rapid development in most UAV techniques has fostered wide attention in the object detection domain.

When focusing on the most common features of an urban environment, buildings constitute the main component of urban areas (Lai et al., 2019). Thus, detection of building

* Corresponding author: Department of Remote Sensing and GIS, Faculty of Geomatics, Sabaragamuwa University of Sri Lanka, Sri Lanka.

E-mail address: r.r.gamage14@gmail.com (R.R. Gamage).

objects from image data is a vital step in most urban related applications. However, objects in UAV images are distributed with great heterogeneity and varying size, and they create great difficulty for extracting buildings using existing methods, which mostly used traditional 2D pixel-based analysis (Zhang et al., 2020). Thus, detection of buildings accurately and automatically has become a hot research topic. To address the aforementioned challenges, this paper presents a new technique which is based on the fusion of point cloud analysis and texture feature analysis, for building detection from UAV images. The texture features can be extracted by using elevation information of each point.

The paper is structured as follows. Section 2 discusses the recent technologies with their significances and limitations. In Section 3, the core method and basic principles of this paper are elaborated in detail. Further, all steps of the method are described. Section 4 describes the experiments that were carried out according to the method, the experimental data, the final results, and the evaluation of the accuracy. Section 5 summarizes the work of this paper and research prospects.

2 Related Work

Building objects are widely used in various applications, including urban planning, cartographic mapping, and land use analysis. Thus, building extraction based image data or laser scanning data is an active research area, and it has been extensively studied for decades in the fields of photogrammetry (Gilani et al., 2016). The developed methods can be broadly classified into three groups based on to data source: (1) 2D image-based methods, (2) 3D point cloud-based methods, and (3) hybrid methods (He et al., 2019).

In recent years, several methods have been developed to extract buildings through 2D imagery. (Ahmadi et al., 2010) proposed an active contour model-based method, while a method based on region growing has been introduced by Ghanea et al. (2014). Further, multiscale morphological index-based method (Huang and Zhang, 2011), object-based method (Chen et al., 2018), and different network-based method (Yang et al., 2018) are most popular methods in the area of building extraction from high resolution image data. Even though these methods have achieved good results, they fail to extract buildings under complex backgrounds of images such as different illumination, occlusion, shadow effect, and geometric deformation. Furthermore, it was difficult to discriminate building from other non-building objects which have similar radiometric signatures. Thus it is concluded that building extraction is difficult by analyzing only spectral information. Therefore, further improvement in building extraction is vital to satisfy various other applications.

Unlike 2D imagery, LiDAR data are more suited for distinguishing building and non-building objects via height variation (Du et al., 2017). Yet, the use of elevation data alone is problematic when separating building and non-building objects with similar heights, for example buildings with trees/smooth canopies surrounding them. Moreover, automatic building extraction is challenging in the contexts of complex shape, occlusion, and size. Therefore, building extraction using a single data type, either 2D images or 3D LiDAR point clouds, remains inadequate (Pirasteh et al., 2019). This problem can be overcome by combining 2D and 3D data (Maltezos and Ioannidis, 2015). Some approaches have been developed to delineate the boundaries of buildings by integrating LiDAR point clouds and orthoimage/image data (Gilani et al., 2016). In general, LiDAR point cloud data are challenging to acquire due to the high cost that is involved. Thus, 2D images and Digital Surface Model (DSM) have been utilized by Tian et al. (2013) for extracting building features.

In recent years, UAV image data have been used widely for various applications as its cost is relatively low with higher flexibility compared with LiDAR (Rosnell and Honkavaara, 2012). Therefore, it is vital to present a method to extract buildings from UAV Image data and image-based derived 3D point clouds. This problem has been solved by Dai et al. (2017) by presenting a method based on RGB- MFV and Support Vector Machine (SVM) classifiers. In this method, buildings were extracted by eliminating vegetation using a certain height threshold. It has shown successful results with simple buildings having linear and perpendicular edges. Unfortunately, this technique was unable to extract buildings when irregular shaped objects exist. Therefore, we present a new method to extract buildings by combining UAV ortho-images, image-derived point clouds, and texture features i.e. elevation map information generated based on the height of each point.

3 Materials and Methodology

3.1 Study Area

The area selected for implementing and checking the accuracy of the proposed method is located in the premises of Sabaragamuwa University of Sri Lanka, approximately at 80.7862^0 N and 6.7157^0 E. Though the selected area does not have the consistency of a very dense urban area, it is a significant built-up area with vegetated areas. Most of the buildings included in this area have simple roofs, and many are restricted to a single floor only, especially the residential buildings. In addition to these, there are faculty buildings, administration offices, quarters, pavilion, and so on and so forth. The total area covered by this test site is around 40 hectares, and consists of more than 100 buildings. The entire working area was divided into six segments, and they were used as the test areas of this research.

3.2 Drone Image Data

The data in this research work were collected on February 14, 2020, from Sabaragamuwa University of Sri Lanka premises in Belihuloya. The Bhoomi Tech (Pvt) Ltd supported the work by acquiring UAV image data. Photographs covering the study area were taken using the phantom 4-DJI UAV, where the on-board camera is equipped with a 1-inch 20-megapixel sensor. A mechanical shutter was used to eliminate the rolling shutter distortion, which can occur when taking images of fast-moving objects or when flying at high speeds. As a result, it can be considered as powerful as many professional cameras. An area of 40 hectares can be covered with approximately 375 images.

3.3 Methodology

The proposed method for building extraction consists of four main stages; (1) Pre-proposing (Point cloud generation), (2) Grid generation, (3) Analysis of height variation and Object point detection, and (4) Building segmentation. The workflow is demonstrated in Fig. 1.

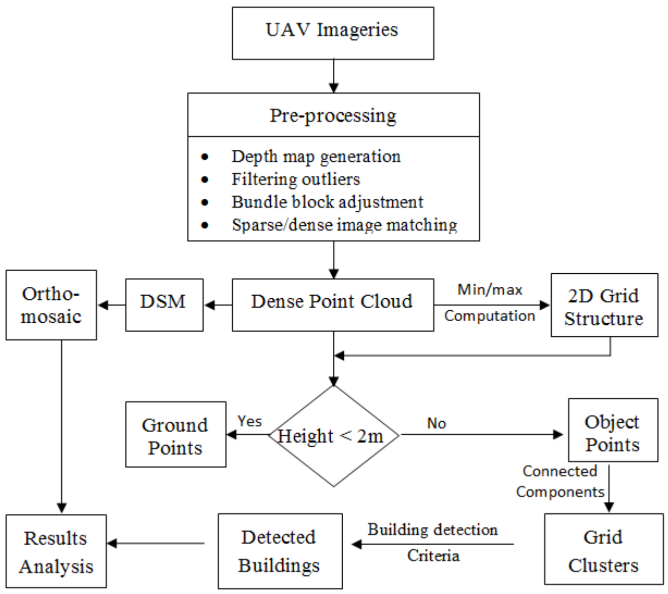

Fig. 1: The utilized framework for building extraction.

3.3.1 Generating point cloud

Generating the point clouds is a pre-processing step of the proposed method. We applied the photogrammetric process to generate a dense point cloud and orthomosaic image from the captured UAV images. For generating the point clouds from UAV images, Agisoft_metashape_professional software was used with high-level setting parameters for each step to preserve the original size of the raw images and also to obtain a more detailed output. At the stage of dense point cloud generation, depth maps were calculated for each image. However, there can be some outliers among the generated points because of noise or badly focused images. In order to solve the problem, a filtering technique can be used. Here, the aggressive depth filtering technique with 'high quality' reconstruction parameter was used for generating dense clouds using the Agisoft software as there were no meaningful small details to be reconstructed. This depth filtering method was reasonable to sort out most of the outliers, and also to generate the most accurate geometry. During the photogrammetric process, bundle block adjustment, and dense image matching techniques were used for generating a photogrammetric point cloud. Bundle adjustment is a photogrammetric technique to combine multiple images of the same scenario into an accurate and efficient 3D reconstruction. It estimates the 3D location of features in the scene, while estimating the camera locations. Dense image matching technique is used for calculating the depth of each pixel. By using these techniques, the software generates the dense point cloud relevant to the UAV images. These generated points have X, Y and Z coordinates relevant to the arbitrary coordinate system. These points also have RGB values. However, the X, Y, and Z coordinates were used in this research.

3.3.2 2D Grid Structure

The irregularly distributed point cloud data which are spread all over the area have three dimensional properties. This point cloud was then converted into a grid structure in order to enhance the speed of data processing for feature extraction. This process was based on the elevation distribution of the point cloud and is easy to operate. First, the grid size was set, based on the average point density of the data set. According to the X and Y coordinates, a grid structure was created for the entire area. This was done by first identifying the maximum and minimum X, Y coordinates in the study area. After that, the grid was divided into segments based on the grid interval that was chosen as a parameter in the proposed methodology. To plot the points into the grid structure, the "mesh' command was used in matlab, and unique "Cell IDs" were created for each grid, based on the position of the grid within the structure. Once the grid structure was created, the generated points were assigned to the respective grid, in accordance with the X, Y coordinates of each point as shown in Fig. 2.

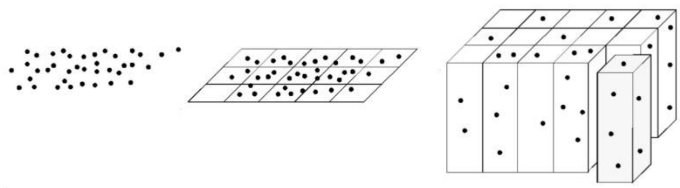

Fig. 2: Analysis of height variation and building point detection Grid structure - (left) Original point cloud in 2D, (middle) 2D grid structure, (right) Grid structure in 3D.

Building points and other points differed from their elevation. Therefore, this height variation was the most useful component in building extraction. With their regular or irregular height variations, building points can be extracted. The process of identifying building areas (point clusters) is performed on grid domain based on this height variation. The presented method is based on the principle that buildings have various properties, for example shape, elevation, unique boundaries, and so on.

The created grids have a different number of points. According to the elevations of the points in a grid, the maximum and the minimum heights were calculated. Further, the point density of each grid was computed. Based on these heights of the grid and point density, each grid was separated as an object or ground grid. Most ground points had approximately similar elevations compared to non-ground points, especially for the local neighborhood. The ground points were also the points with lowest elevation when compared to other points. Similarly, ground has a higher reflection than non-ground points. When a unique area of ground and non-ground is compared, the point density of ground area is much higher than non-ground area. After assigning grids into object and non-objects groups, the grids assigned as objects were clustered based on the connected component analysis. These clusters, called the “grid group”, were used to further analyze grids. Each grid group was composed of several grids that were adjacent to each other, and that had the same geometric properties. Before clustering the objects, certain grids which disagreed with the threshold of height variation were removed. For that, a 3*3 moving window was used. Then the rest of the grids were clustered in accordance to their geometric properties.

The next step of the methodology was to classify grids belongs to building points. For this, different building detection criteria were introduced to separate building clusters from other object clusters, based on the geometric properties and grid adjacency. These criteria definitions were based on the assumption that the adjacent grids with small elevation differences are more likely belong to the same object. This assumption is derived from the fact that the object has continuity in horizontal direction. These criteria are mostly based on the area, perimeter/perimeter length, height and so on. The complete workflow of this part is shown in Fig. 3. It consists of several steps which improve the quality of the outputs.

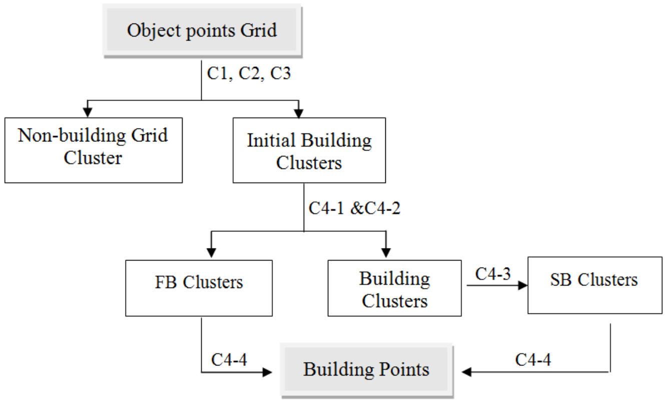

Fig. 3: Workflow of non-ground classification.

Table 1: Building detection criteria

Criteria	Function	Outputs
1. Area analysis	1. Area > T1	Identify initial building clusters
2. Height & Area analysis	1. No. of grids (h_{max} > 2m) >T1	Remove connected non-building grids and non-building clusters
3. Shape-based analysis	1. LW-min < length - width ratio < LW-max 2. circularity (Circularity = (4 * pi * Area) / Perimeter^2) >1 (Takashimizu and Iiyoshi, 2016)	Remove the irregular or linear shaped objects such as trees, flower fences & etc.
4. Geometric properties and grid adjacency analysis	1. Height variation $Diff_{(hmax-hmin)} < Bh$ $\left(\sum_{i=0}^{n-1} hmax_i - \sum_{j=1}^{n} hmax_{ij}\right) < 0.5$ 2. Plane surface $No.of Grid[(\sum_{i=0}^{n-1} hmax_i - \sum_{j=1}^{n} hmax_{ij}) < 0.5)] > \text{Tmin}$ $(Avg\left(\sum_{i=0}^{n} h_{max}\right) < 0.5$ 3. Direction (slope) of orientation is unique (compute gradient – check first order derivatives) 4. Point density variation is unique	Remove tree clusters and correctly identify Building clusters

Where;

T1 – minimum building size; Bh – building height ; LW-min = minimum building length-width ratio; LW-max = maximum building length-width ratio; T2 = no of grids of smallest plane surface.

As the first step, according to the cluster ID, the area was calculated. This first criterion helps to identify initial building clusters. However, the clusters belonging to vegetation can be mis-classified as building clusters due to their larger size. Thus, the challenge was to separate the building points from other non-ground points, for example tree points. Those clusters were removed by introducing a shape-based analysis method. Under this step, two main criteria were defined: (1) length - width ratio and (2) circularity. Shape of each cluster was identified, and clusters were divided into three main groups as linear, planar and spherical. This shape analysis method isolated the building clusters successful by removing most of the irregular or linear shaped objects such as trees, flower beds, fences and so on, successfully. However, all points related to the trees were difficult to identify when the size and shape of clusters are almost similar to the buildings criteria. Furthermore, vegetation surrounding the buildings with similar heights was recognized as parts of buildings.

In order to improve the correctness of the selected building from the above conditions, geometric properties of buildings were further included to this process. Here, it was assumed that most buildings have planar surfaces. Not only that, but their height variation and the direction (slope) of orientation are also unique. In tree clusters, the maximum height can be seen in the middle of the cluster, and their point density is also higher than the outside pixels. Therefore, the next criteria were defined based on these geometric properties of buildings in order to demarcate buildings accurately from other objects. All criteria are listed in the Table 1.

Based on these conditions, the clusters which satisfied the above conditions were selected as building clusters. Then the grids that were within those clusters were selected, and the points in those grids were selected as building points. Using selected building points, building boundaries were detected based on the canny filtering technique. The canny filtering for edge detection is one standard technique mostly used to identify edges/boundaries of objects from 2D images due to its better localization and good recall. Since the outcome of the building point extraction was converted to raster data, canny filtering was applied straightaway. Therefore, in the first phase, grid data was converted into 2D image and edges were extracted using the Canny edge detector technique. Then, 2D edge pixels were converted into 3D point cloud in order to identify 3D edges.

4 Results and Discussion

The study area was divided into six test sites, to cover all the buildings. Given that the total area was quite large, the number of points was massive. In general, LiDAR points create only roofs of buildings and vegetation as a canopy type, but the points generated under this method were different from LiDAR points. Building's façade was also generated as points in the overlapped area. Furthermore, under-the-trees points were generated through the vegetation. A dense point cloud, having 53,958,325 points, was generated covering the entire research area of 725,058.6 m^2. The point density of the generated point cloud was 75 points per m^2. The orthomosaic was also generated according to the work flow of Agisoft software using the generated dense point cloud (see Fig. 4). The pixel size of the orthophoto was 3.97 cm per pixel.

Fig. 4: Generated orthomosaic image

The full area was first divided into small grids, and a 2D grid structure was developed in order to divide points obtained from the thick point generation process into different categories. A 2D grid network was used for analyzing the denser point cloud data as a data structure for the identification of objects, especially buildings. The results, however, depend on the size of the grid. Usually by adding points of neighboring objects, a large size may be misclassified. Therefore, it is possible to pick small size gaps with grids without any lines. As such, by looking at the thick point clouds and its estimated resolution, the grid interval was chosen. Using the created 2D grid structure, points were classified into two major groups as ground and non-ground. Based on their height differences, the selected non-ground pixels (points) were grouped into clusters.

According to the given parameters, buildings were categorized in 2D grids. The extracted buildings from each area are visualized in Fig. 5. Blue, orange and white colour represent obtained building points, non-building points, and areas where the number of points is zero, respectively. In order to obtain best results, the parameter tuning process was adopted.

4.1 Building Extraction Performance

In order to prove the effectiveness of the proposed method, the experimental results were compared with reference data. The reference data was generated using the manual extraction process with visual interpretation of orthophotos. Based on the study area, different parameters were used for each step (Table 2).

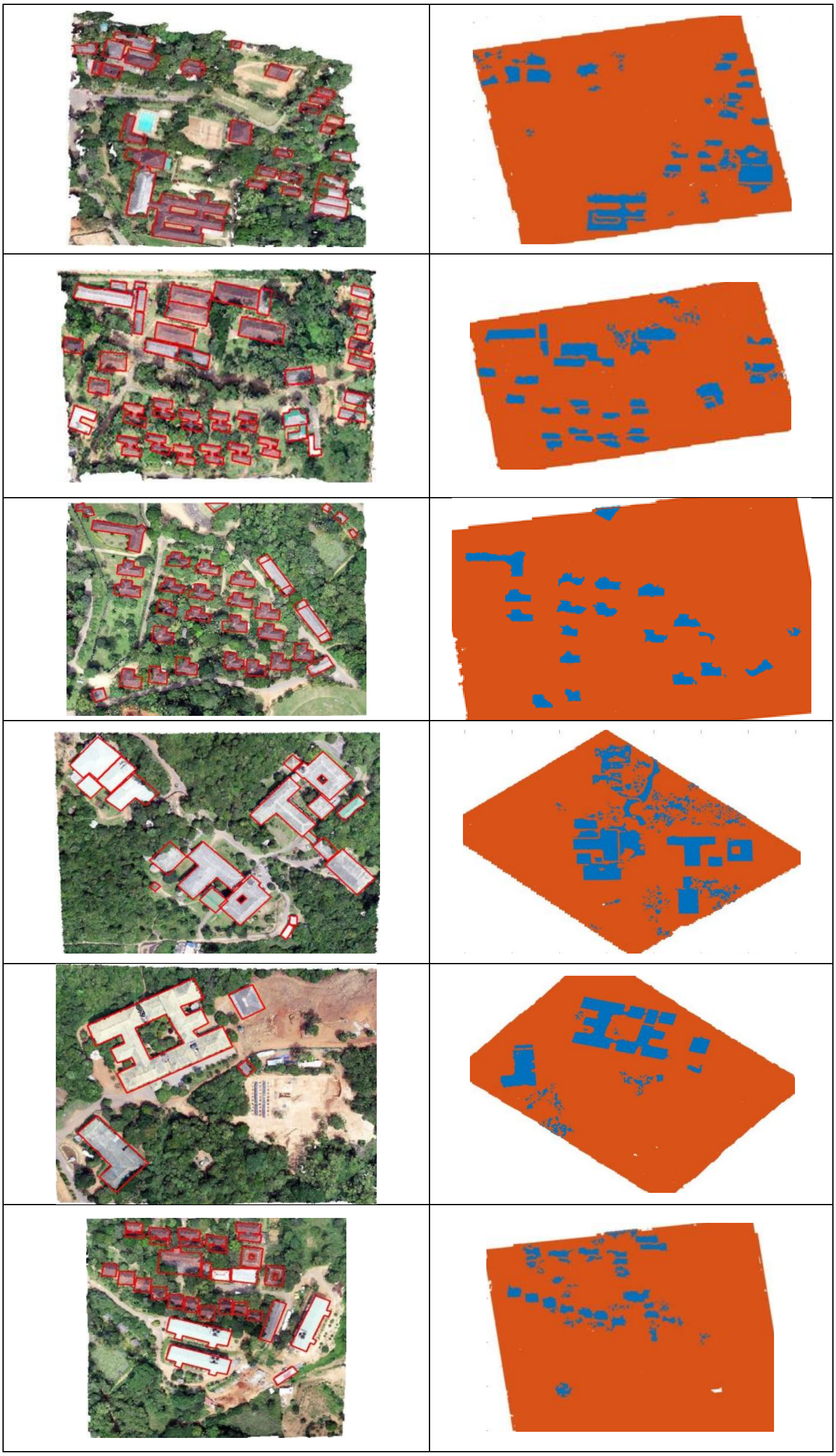

Fig. 5: Extracted buildings: Orthomosaic images (1st column) and resulted building points (2nd column) on SA 1, SA 2, SA 3, SA 4, SA 5, & SA 6 (from top to bottom respectively)

Table 2: Parameter setting for each study area

Study area / Parameters	SA 01	SA 02	SA 03	SA 04	SA 05	SA 06
Grid Interval	0.5	0.5	0.5	0.5	0.5	0.5
Average Ground Height	10	15	18	18	2	15
No. of points in a cell	20	20	20	20	20	20
Slope Height Range (Ground)	0.45	0.4	0.4	0.4	0.37	0.4
Minimum Area of Building	42.136	100	40.658	100	62.755	100
Maximum Area of Building	3635	968.219	583.598	3196	4256.7	1500
Slope Height Range (Building)	0.25	0.2	0.25	0.25	0.19	0.17

There was a deviation in the number of included buildings and the number of extracted buildings. Extracted buildings were accurately extracted, but due to the dense vegetation and undulation of the surface, some buildings were not extracted. The number of the un-extracted building increased, especially when the area was highly undulated with sudden changes in elevations. The performance analysis of the building extraction is presented in Fig. 6.

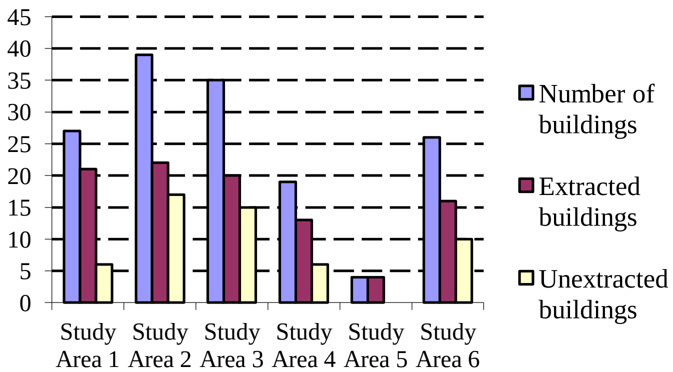

Fig. 6: Quantitative results of the extracted buildings of each test area.

The results of building extraction using the proposed method were evaluated by overlapping the results with the orthophotos. The most popular metrics to assess the results of building detection are Completeness (COMP), Correctness (CORR), and Quality (QUAL). Four indicators were used to evaluate the performance of extracted buildings based on the above three matrices. They are (1) the number of buildings correctly classified as buildings (TP), (2) the number of non-buildings incorrectly classified as buildings (FP), (3) the number of non-buildings correctly classified as non-buildings (TN), and (4) the number of buildings incorrectly classified as non-buildings (FN).

Three metrics: (1) completeness, (2) correctness, and (3) quality, were computed to evaluate the results of the proposed method using the following equations (1), (2), and (3), respectively (Awrangjeb and Fraser, 2014);

$$\text{Completeness (COMP)} = TP / (TP + FN) \tag{1}$$

$$\text{Correctness (CORR)} = TP / (TP + FP) \tag{2}$$

$$\text{Quality (QUAL)} = TP / (TP + FN + FP) \tag{3}$$

The evaluation results based on the above three indices are shown in Table 3. It shows that the proposed method has been able to achieve state-of-the –art results in most of the test areas, except for test area SA4. According to Table 3, it can be concluded that the SA3 has the best results compared to the other areas. This can be clarified by the lower occlusion effect/data gaps in this area. However, SA1, SA2, SA5 and SA6 are also showing good results, by having more than 60 % of quality. It also shows that the proposed method is rather reliable for detecting buildings, with less than 10 FP value for each dataset. The main reason is the largest tree clusters that can be seen in the study area and the complexity. However, it does not influence for the overall figure of correctness (see Table 3). These quality values disclose the robustness of the shape and height analysis criteria.

In contrast, less than 20% of the buildings were not detected, except for the study area SA4 which has a 30% FN value. One reason for this FN error is gaps in the point clouds, i.e., an incomplete building area.

Table 3: Overall statistics of extracted buildings in test areas SA1 – SA 6

Study Area	Number of total buildings	TP	FP	FN	COMP	CORR	QUAL
SA1	27	21	4	6	77.78%	80.76%	67.74%
SA2	39	30	6	9	76.92%	83.33%	66.66%
SA3	35	30	7	5	85.71%	81.08%	71.42%
SA4	19	13	5	6	68.42%	72.22%	54.16%
SA5	4	4	2	0	100%	66.66%	66.66%
SA6	26	21	6	5	80.76%	77.77%	65.66%

Where: TP – number of buildings correctly classified as buildings, FP – number of non-buildings incorrectly classified as buildings, FN – number of buildings incorrectly classified as non-buildings, COMP – Completeness, CORR – Correctness, and QUAL – Quality.

Differences between the obtained results and the actual ground were observed (see Fig. 7). In most cases, points belonging to natural features, in particular points on trees and flat surfaces of the land, were correctly identified. Many tree canopies tend to be like a building's roof, especially tree canopies on a hill, with variations in height in the landscape. Tree canopies parallel to the roof were often interpreted as a building level. Trees that were mostly higher than buildings and situated close to the building, were attached to the building roof. Some flat ground zones were also categorized as smooth and less undulating building points.

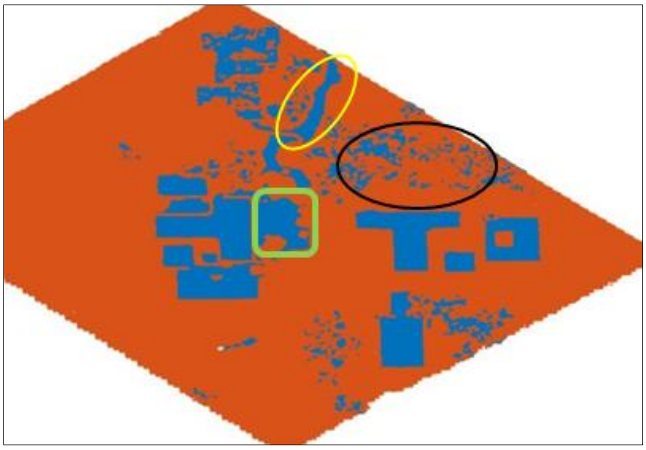

Fig. 7: Deviations of obtained result (Black-Points on trees, Yellow-Points on a road, Green-Points on a flat surface)

To remove these types of errors, the conditions used to remove buildings in the point cloud must be improved. Additionally, other criteria can be added to eliminate trees, so that the misclassification of canopies as building points is easier to reduce. The average elevation of the area of study was about 50 m. As the used classification criteria depend on the elevation, this flat surface is mainly classified as building points.

In less undulated areas, it is possible to achieve more precise results based on the proposed method. As the field site included a highly undulating low jungle, advanced conditions are needed to obtain more detailed results during building extraction.

According to the conditions used, the algorithm was able to extract a higher number of buildings. Each study area was stored as a .txt file, and the size of a file was about 300MB. As a consequence, the algorithm took 13-15 hours to complete the process. Because of that, six machines were used to process one round of samples for each study area. Hence, identifying changes in the images when the parameters were changed, consumed a significant time.

5 Conclusion

Extracting building points from low-cost UAV imageries is a revolutionary idea for most. This is because the present society wastes a lot of money to detect features and acquire their boundaries in a legal background. UAV based mapping is developing fast to assist resolve these problems. Not only that, UAV is the cheapest and fastest way of mapping in the world.

Using low-cost UAV imageries and generating dense cloud is a cheaper way for feature extraction, boundary detection, as well as in many other applications. Using the LiDAR point cloud is more costly when compared to using UAVs, and the latter can do the job in a different way. As a developing country, this method is more suitable to be used for surveying purposes.

Acknowledgement

The authors would like to thank Bhoomi Tech (Pvt) Ltd for their support on acquisition of UAV data.

Author Contributions

Conceptualization, methodology, analysis, writing—original draft preparation, by R.R.G., and writing—review and editing, by H.A.N. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

Ahmadi, S., Zoej, M. V., Ebadi, H., Moghaddam, H. A., Mohammadzadeh, A., 2010. Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours. International Journal of Applied Earth Observation and Geoinformation. 12, 150-157.

Awrangjeb, M., Fraser, C. S., 2014. An automatic and threshold-free performance evaluation system for building extraction techniques from airborne LIDAR data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 7, 4184-4198.

Chen, R., Li, X., Li, J., 2018. Object-based features for house detection from RGB high-resolution images. Remote Sensing. 10, 451.

Dai, Y., Gong, J., Li, Y., Feng, Q., 2017. Building segmentation and outline extraction from UAV image-derived point clouds by a line growing algorithm. International Journal of Digital Earth. 10, 1077-1097.

Du, S., Zhang, Y., Zou, Z., Xu, S., He, X., Chen, S., 2017. Automatic building extraction from LiDAR data fusion of point and grid-based features. ISPRS journal of photogrammetry and remote sensing. 130, 294-307.

Gerke, M., Przybilla, H.-J., 2016. Accuracy analysis of photogrammetric UAV image blocks: Influence of onboard RTK-GNSS and cross flight patterns. Photogrammetrie, Fernerkundung, Geoinformation (PFG). 17-30.

Ghanea, M., Moallem, P., Momeni, M., 2014. Automatic building extraction in dense urban areas through GeoEye multispectral imagery. International journal of remote sensing. 35, 5094-5119.

Gilani, S. A. N., Awrangjeb, M., Lu, G., 2016. An automatic building extraction and regularisation technique using lidar point cloud data and orthoimage. Remote Sensing. 8, 258.

Harwin, S., Lucieer, A., 2012. Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. Remote Sensing. 4, 1573-1599.

He, H., Zhou, J., Chen, M., Chen, T., Li, D., Cheng, P., 2019. Building extraction from UAV images jointly using 6D-SLIC and multiscale Siamese convolutional networks. Remote Sensing. 11, 1040.

Huang, X., Zhang, L., 2011. A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery. Photogrammetric Engineering & Remote Sensing. 77, 721-732.

Lai, X., Yang, J., Li, Y., Wang, M., 2019. A building extraction approach based on the fusion of LiDAR point cloud and elevation map texture features. Remote Sensing. 11, 1636.

Maltezos, E., Ioannidis, C., 2015. Automatic Detection of Building Points from LIDAR and Dense Image Matching Point Clouds. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences. 2.

Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: a review. Applied geomatics. 6, 1-15.

Pirasteh, S., Rashidi, P., Rastiveis, H., Huang, S., Zhu, Q., Liu, G., Li, Y., Li, J., Seydipour, E., 2019. Developing an algorithm for buildings extraction and determining changes from airborne LiDAR, and comparing with R-CNN method from drone images. Remote Sensing. 11, 1272.

Ramon Soria, P., Bevec, R., Arrue, B. C., Ude, A., Ollero, A., 2016. Extracting objects for aerial manipulation on UAVs using low cost stereo sensors. Sensors. 16, 700.

Rosnell, T., Honkavaara, E., 2012. Point cloud generation from aerial image data acquired by a quadrocopter type micro unmanned aerial vehicle and a digital still camera. Sensors. 12, 453-480.

Takashimizu, Y., Iiyoshi, M., 2016. New parameter of roundness R: circularity corrected by aspect ratio. Progress in Earth and Planetary Science. 3, 1-16.

Tian, J., Cui, S., Reinartz, P., 2013. Building change detection based on satellite stereo imagery and digital surface models. IEEE Transactions on Geoscience and Remote Sensing. 52, 406-417.

Yang, H., Wu, P., Yao, X., Wu, Y., Wang, B., Xu, Y., 2018. Building extraction in very high resolution imagery by dense-attention networks. Remote Sensing. 10, 1768.

Yao, H., Qin, R., Chen, X., 2019. Unmanned aerial vehicle for remote sensing applications—A review. Remote Sensing. 11, 1443.

Zhang, R., Shao, Z., Huang, X., Wang, J., Li, D., 2020. Object detection in UAV images via global density fused convolutional network. Remote Sensing. 12, 3140.