AUTOMATIC GEOREFERNCING OF CLOSE-RANGE FAÇADE IMAGES ACQUIRED IN AN NARROW AND LONG ALLEYWAY USING RTK DRONE IMAGES

K. Park[1], S. Ham[1], I. Lee[1] *

[1] Dept. of Geoinformatics, University of Seoul – (minpkang, flyhamsw)@gmail.com, iplee@uos.ac.kr

Commission II, WG II/1

KEY WORDS: Georeferencing, RTK drone, Action camera, Close-range images, Alleyway, 3D modelling

ABSTRACT :

The city of Seoul has selected Sewoon market building and its surrounding district as part of the urban regeneration zone, and currently has been promoting the project. To monitor results of the project regularly, the city has been trying to utilize a 3 dimension model of the area. In the case of buildings placed in narrow alleyways in the district, however, it is limited to generate 3D model of the buildings due to some factors. Therefore, in this study, a 3D model of façade of the building was created, using a RTK drone and action camera only. First method is to estimate of location of conjugate points using Structure from Motion, after setting conjugate points between images of the drone. Second method is to georeference action camera images by setting drone images as the reference images itself without the process of estimating location of the conjugate points. As a result of preliminary experiments to verify the two methods, the error of each method did not exceed a maximum of 0.030m. Based on the result, we created 3D models of façade of the building in the alleyway, which is located at the intersection of Donhwamoon-ro 2 gil and Jong-ro 24 gil, and calculated absolute distance between the models. And the comparison showed that the difference was about 0.010m on average.

1. INTRODUCTION

Urban regeneration is an urban policy that attempts to solve problems of low-income neighbourhood by utilizing the historical, cultural characteristics of the city. The city of Seoul, Korea, selected 13 areas as 'Urban regeneration active area', and has implemented the project since 2015. One of the areas is Sewoon market and its surroundings, which are located at Jangsa-dong, Jongno-gu. The district was slum in the 1960s, and Sewoon market was built for environmental improvement and promoting the growth of the city in 1966. At the time of its establishment, Sewoon market was Korea's first mixed-use apartment building called as 'Sewoon, City in the city', and the surrounding area became the main central commercial district of Seoul thanks to the Sewoon market. However, the development of Gangnam area had led to a rapid decline of the district since 1970s. Since then, the Sewoon market, the commercial centre, and its surroundings declined, and discussions on redevelopment continued. But it could not be implemented, however, due to cost and conflict with residents, etc., the area has fell into a representative low-income neighbourhood of the Seoul. And then finally, the Sewoon market and its surroundings was selected as 'Urban regeneration active area' in 2015.

In the project, the city of Seoul has utilized a three-dimension model of the Sewoon market area for monitoring results of the project. For generating the 3D model, aerial images and construction drawings have been used. With this model, the Seoul has tried to monitor by utilizing a similar approach as "Digital Twin", which combines real-time data of the building obtained by various sensors on the model. However, this kind of attempts are relatively difficult for buildings in alleyways around the Sewoon market. This is

because it is very hard to acquire data for creating 3D model. It is due to the fact that the buildings are located at narrow and long alleys and there are relatively many floating populations. This study aimed to generate 3D model of façade of buildings which are in the environment. For that, only RTK drone and action camera were used without a total station.
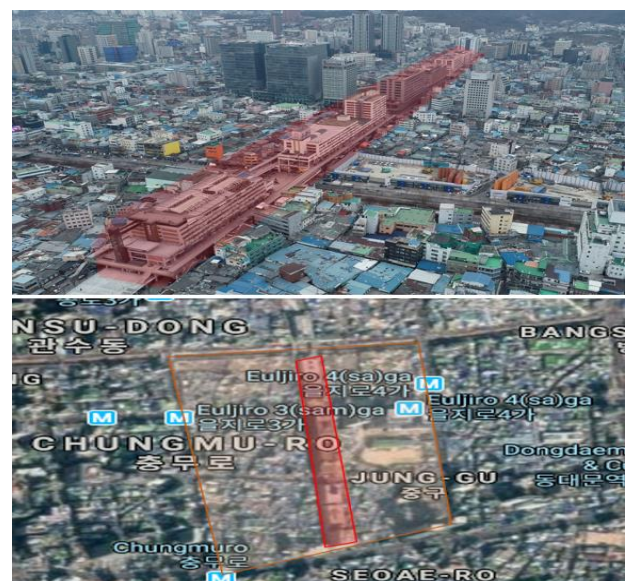

Figure 1. The Sewoon market building and its surroundings. the red square is Sewoon market building and the orange is its surroundings.

---

* Corresponding author

2. ACQUIRING AND PROCESSING DATA

2.1 RTK drone and Action camera

RTK provides very precise location data based on GNSS. So, 3D models using RTK drone can be expected to have high precision (Tomaštík et al., 2019), and currently it is easy to find study cases of creating 3D models by the RTK drone (Urban et al., 2019, Taddia et al., 2019). In this research, however, there are limitations to generate the model only using the drone. This is due to the environment in which building are laid and the structural features of the building. In the alleyways around Sewoon market, lots of telegraph poles are distributed around buildings which usually lower than the poles. As a result, a landscape with multiple electric wires is formed on the alleyway, which not only makes it hard to take complete images of the buildings using the drone, but even makes close-ups difficult. Besides, shading devices attached to façades also act as barrier to modelling using the drone. Sections without the electric wires could be rarely found, but occlusion areas usually occur in most images due to the position of the drone and the sunshades. Modelling using RTK drone is clearly an efficient method, but it would be not possible to use it in the research environment, considering these difficulties.

Figure 2. 4 pictures are the view of the alleyway. As you can see in pictures, the width of the road is very narrow. It is very easy to find telegraph poles, power lines and shading devices in the alley.

Action cameras are tools characterized by high portability, and these are getting attention as tools for personal media, leisure activities recently. These provide various shooting methods such as time-lapse, and it is easy to find products released in wide-angle. The width of the alleyways in this study is about 2 meters wide, and there are many smaller areas. In such an environment, it is difficult to efficiently take images of façades located on the side of alleyways with a regular camera. However, by using a wide-angle action camera, not only is it easy to overcome the limitation, but also the time-lapse function can be used to acquire data more easily. It is indispensable, however, to acquire ground control points separately for creating 3D model with images of the camera. It is common to use a total station to acquire ground control points, but it cannot help but being limited when considering the environment. This is because the width of the alleyway is too narrow to set up a total station, and it is difficult to occupy a space stably for measurement because of the floating

population. In addition, even if the measurement is performed by occupying a specific space stably, it is necessary to continuously change the position of total station to secure the field of view.

2.2 Data processing methods

In summary, in the case of RTK drone images, it is possible to secure the very precise pose and location of images, but it is very hard to obtain complete façade images of buildings due to the environment and the characteristics of the target. On the other hand, in the case of action camera images, it is possible to obtain complete images of the facades efficiently, but it is essential to take ground control points separately for generating 3D model using a total station. However, as mentioned, the environment limits the acquisition of ground control points using a total station. Therefore, in this research, the action camera images were georeferenced using the images taken with the RTK drone, which can secure a high location precision, instead of total station, as the reference data, and generated 3D model of the façade. For this, the areas where the occlusion region in the RTK drone images did not occur were taken overlapping with an action camera, and then the data were processed in two ways.

The first is a method(A) of estimating location of conjugate points between RTK drone images, and these estimated conjugate points are used as ground control points. After selecting conjugate points between images, it is possible to estimate precise location values of the points using the Structure from Motion algorithm(SfM). And then, a 3D model is able to be generated after georeferencing the action camera images using the estimated points as ground control points. For this, the conjugate points must be extracted from overlapping area with the action camera images. The second is a method(B) of performing the SfM with the both types of images at once by setting the RTK drone images as the reference data itself. For setting the drone images as the reference data, a relatively large weight should be given to the location data of the RTK drone images to narrow the adjustment range in the SfM. It is different from the former method in that ground control points are not estimated from the drone images separately, and because of it, the second is relatively more efficient. But, if there is a large difference in the scale between the two types of images due to some factors like the difference in elevation, automatic matching between feature points may not work properly, so, a process of manually selecting conjugate points between the images could be necessary.
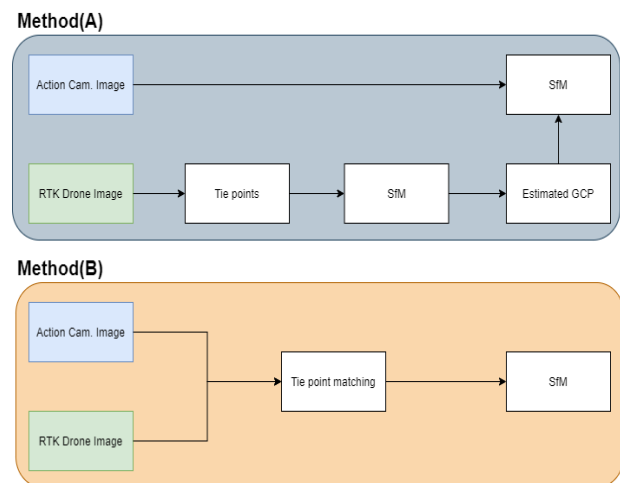

Figure 3. Workflow of method (A) and (B).

Before applying these methods, the process of estimating and adjusting the interior parameters of the drone and action camera should be preceded. The methods commonly estimate the external parameters of the action camera using the drone images. Since the interior parameters are closely related to the exterior parameters, it is essential to precisely estimate the interior parameters in advance and apply them consistently. The interior parameters of the drone and action camera used in this study were estimated in the preliminary experiments. For this, self-calibration was performed using ground control points acquired by a total station.

3. METHOD VERIFICATION AND APPLICATION

3.1 Method verification

Before applying the two methods suggested in this research to the site, preliminary experiments were conducted on the Jeon-nong Hall, a building in the University of Seoul, Korea, to verify the accuracy of the both methods. For the verification, ground control points on façade of the building were acquired using the total station, and the preliminary experiments were conducted in two ways. First, root mean square error(RMSE) was calculated by using the ground control points as check points of 3D models made by the two methods suggested in this study. Also, using the ground control points, the pose and location values of each image acquired by the action camera were estimated through the SfM. And then, the estimated values of the images were compared with the pose and location of the same action camera images georeferenced by the methods presented in this study, and used the difference as second index to judge the accuracy of each method. In this research, a Phantom 4 RTK drone(P4RTK) and Osmo Action camera(OSMO) from DJI were used. In the case of Osmo Action, pose and location data of images are not tagged. For total station, TOPCON OS-102 CS0133(OSMO) was used, and for processing data, Photoscan pro from Agi and CloudCompare were used.

The southern façade of the Jeon-nong hall is about 10m high and about 20m wide. Firstly, data to verify the two methods presented in this study were generated, and the data was assumed as the true data. For this, a total of 136 façade images was acquired with OSMO, and 18 ground control points (EPSG5186, GRS80) were acquired from the southern façade with TOPCON. And then, the SfM was performed using the ground control points and images, through which the pose and location values of each image were estimated. To this end, 9 of the 18 points were used as control points, and the remaining 9 as check points. As a result of the SfM, the RMSE of the control points in easting, northing, altitude was 0.007m, 0.005m, and 0.013m, and the RMSE of the check points were 0.005m, 0.014m, 0.007m. In addition, the interior parameters of P4RTK and OSMO were self-calibrated using the ground control points acquired in this process, and then the parameters were used for all preliminary experiments and field applications.

	Pixel size	Focal length	Cx	Cy
OSMO	0.0015345	1716.0924	2.7461	-18.005
P4RTK	0.0024123	3672.4596	20.069	8.8686

Table 1. Self-calibrated interior orientation parameters of OSMO and P4RTK by GCPs. The unit of pixel is mm and the others are pixels.

Subsequently, two 3D models of the same southern façade were made by the two methods suggested in this study. First, in order to acquire data to be used as the reference data in each method, 10 vertical images were acquired by flying P4RTK from the south to the north at an altitude of 15m. Because it is hard to take a lot of images by the drones in the alleyway of Sewoon market, the original experiment area, only 10 images were taken intentionally in the preliminary experiment.
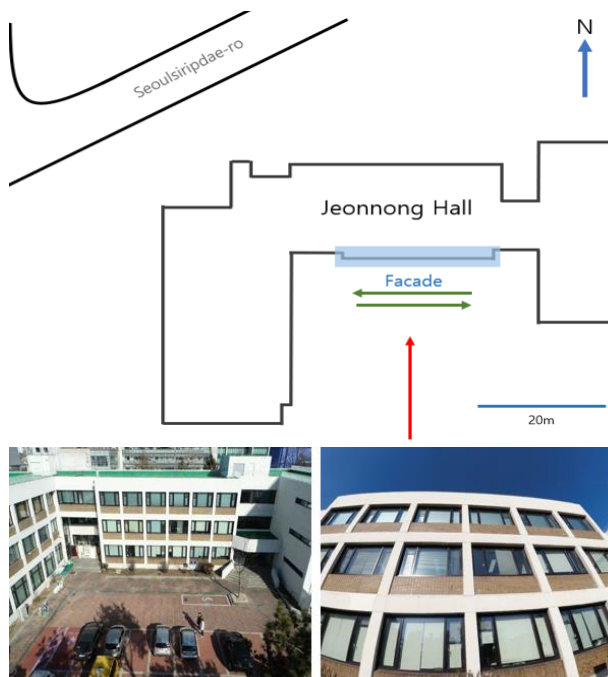

Figure 4. Red arrow shows how RTK drone moves from the south to the north, and the lower left is an image taken by the drone. Green arrows are the action camera flow line. The lower right is an image taken by the action camera.

The method(A) is to estimate the coordinates of the conjugate points through the SfM between the drone images, and for this, a total of 16 conjugate points was set. When setting the conjugate points, these were not overlapped with the ground control points acquired by the total station. Of the conjugate points, 9 points were used as ground control points for the action camera images, and 6 points acquired by the total station were used as check points. The action camera images used at this step were identical with the images used for making the true data. The RMSE of the control points in easting, northing, altitude was 0.008m, 0.008m, 0.006m, and the RMSE of the check points was 0.060m, 0.033m, 0.042m respectively. Next, through above process, the pose and location data of the action camera images were compared with the action camera images georeferenced with the total station.

As a result of comparing the pose and location value of each image, the difference of easting, northing, altitude were -0.093m, -0.024m, 0.037m, and roll, pitch, yaw were -0.253°, -0.131°, -0.125° on average. As it shows, it can be confirmed that the georeferenced pose and location of images through method(A) were very similar with the values georeferenced by the ground control points acquired by the total station. Based on the results, it could be concluded that the 3D model generated by method(A) also shows an acceptable level of accuracy.

RMSE	Method(A)	Average Error	Method(A)
Northing	+0.059m	Northing	-0.093m
Easting	+0.033m	Easting	-0.024m
Altitude	+0.042m	Altitude	+0.037m
		Roll	-0.253°-
		Pitch	-0.131°
		Yaw	0.125°

Table 2. RMSE of the check points and positional and orientational error of the images generated by Method(A).
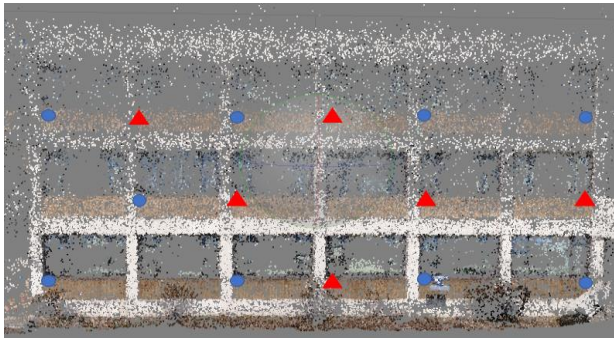

Figure 5. The red triangles are check points from the total-station, and the blue dots are the estimated GCP from method(A).

The method(B) is to use the RTK drone images itself as the reference images for georeferencing the action camera images without the process of extracting ground control points from drone images like method(A). For this, both of images from the drone and action camera are put together in the SfM and processed. At this step, by setting relatively high accuracy weight for the pose and location value of the drone images, it is necessary for fixing the values to be not changed in the SfM. In this study, the weight of the drone images(0.1m) was set to 100 times the weight of the action camera images(10m). Even if the two images are processed together, if the weights are not set as above, the values of the drone images should be changed due to an error in the action camera images, so the drone images cannot function as the reference images.

To verify the accuracy of the method(B), 7 ground control points acquired by the total station were used as check points as did in method(A). As a result, the RMSE of easting, northing, altitude was 0.020m, 0.031m, 0.022m, respectively. Compared to the estimated pose and location of the action camera images based on the ground control points, the average difference of easting, northing, altitude was -0.011m, 0.033m, -0.020m, and the average error of roll, pitch, yaw was 0.004°, -0.388°, -0.004°. In the case of this method as well, it was confirmed that the results were very similar to those obtained by the ground control points acquired by the total station.

RMSE	Method(B)	Average Error	Method(B)
Northing	+0.020m	Northing	-0.011m
Easting	+0.031m	Easting	+0.033m
Altitude	+0.022m	Altitude	-0.020m
		Roll	+0.004°
		Pitch	-0.388°
		Yaw	-0.004°

Table 3. RMSE of the check points and positional and orientational error of the images generated by Method(A).
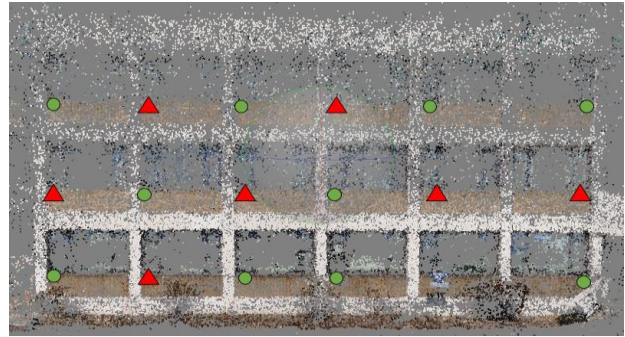

Figure 6. Red triangles are check points from the total-station, and Green dots are tie points between pictures of RTK drone and action camera.

In conclusion, the both of methods showed similar results with the classical photogrammetric method using the total station. Although the method(A) showed a maximum error of about 0.059m from northing, and method(B) showed 0.030m from easting, it was judged as an acceptable RMSE error. When comparing the georeferenced pose and location values, it can be seen that both methods have estimated similar pose and location values with the classical method on average.

3.2 Field application

Among the alleyways around Sewoon market, the façade of a building located at the intersection of Donhwamoon-ro 2 gil and Jong-ro 24 gil alleyway was selected. A drone and action camera were used in the field application, which were identical with ones used in the preliminary experiments. Using the drone, a total of 9 vertical images were taken at an altitude of about 40m while flying the drone from the east to the west in the direction of looking at the façade of the building. Along with the drone, a total of 296 façade images was taken using the action camera, and those images included a region overlapping with the drone images. In the drone images, it is difficult to see the bottom of the building in detail because of the occlusion area caused by the shading devices, and the width of the alleyway, signboard, telephone pole. Therefore, as taking images of upper part of the building by the action camera intentionally, overlapping areas between the drone and action camera images were created.

Figure 7. the green line and red line are Jong-ro 24 gil and Donhwamoon-ro 2 gil respectively. The red box is the target façade of building. The Sewoon market is on the right side of the image.

Figure 8. Picture (a) is one of images taken by the drone flying west to east. Others were taken with the action camera.

In order to apply the method(A), a total of 10 conjugate points was set in the overlapping area with the action camera in the image taken with the drone, and then the SfM is used to estimated location of the conjugate points. Of the 10 estimated coordinates, 5 points were used as control points and the remaining 5 as check points.

	Easting(m)	Northing(m)	Altitude
Point 1	199472.3054	552198.4423	47.3807
Point 2	199472.5342	552200.4637	47.3921
Point 3	199473.0634	552205.2988	47.2872
Point 4	199473.9136	552210.5922	47.5024
Point 5	199475.2853	552219.0198	48.1757
Point 6	199474.4733	552213.4384	47.2520
Point 7	199475.3458	552220.583	45.5020
Point 8	199472.8718	552207.9859	45.8495
Point 9	199472.326	552203.0655	45.8779
Point 10	199471.9246	552198.1671	46.2351

Table 4. Estimated coordinates of tie points on method(A)

As a result of georeferencing action camera images using the control points, the easting, northing, altitude RMSE of the control points were 0.006m, 0.007m, 0.017m, respectively, and the RMSE of the check points were 0.008m, 0.012m, 0.019m.

In the case of method(B), due to the altitude of the drone, there was a large difference in scale between the drone and camera images, which caused a problem in the process of automatic feature matching between the images. To solve the problem, as mentioned above, conjugate points were manually set between two types of images, which were at same location with the conjugate points used in the method(A). In the preliminary experiments, using the ground control points acquired by the total station, the accuracy of the method(B) could be verified. However, in field application, because it was impossible to use the total station and the SfM was performed by setting only the conjugate points, there was a problem in that the RMSE of the 3D model made by the method(B) could not be calculated. Therefore, the similarity between the two 3D models was confirmed using cloud-to-cloud distance method(C2C) between the dense point cloud models of the method(A) and (B). In the method(A), since the RMSE could be calculated, it was judged that the accuracy of the method(B) model could be indirectly verified by comparing C2C with the 3D model of the method(A). As a result of comparing C2C,

there was an average difference of 0.011m between the two models, and the standard deviation was 0.011m.
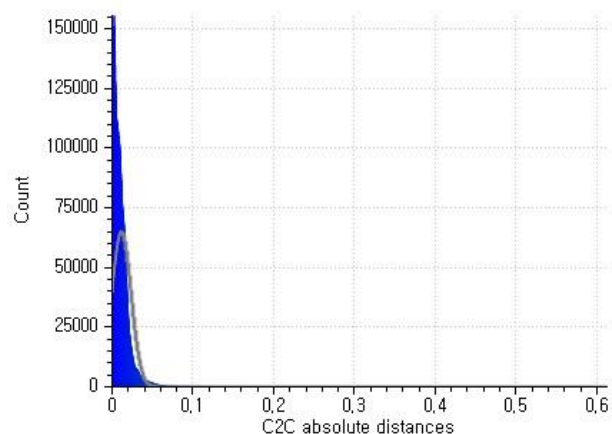

Figure 9. There are the figures of the 3D model of the façade based on method(A) and (B) from the top. Third and fourth figures are the result of C2C.

4. CONCLUSION

This study was conducted for the purpose of creating a 3D model for monitoring the urban regeneration projects. However, the target building was located in a narrow and long alleyway, and there were limitations in the modelling method using a total station due to obstacles such as telegraph poles. Therefore, in this study, two kinds of methods using complimentarily RTK drone and action camera are proposed. Not only are the two methods relatively flexible to environmental constraint, but also ensure efficiency in acquiring and processing data. In

addition, it was confirmed from the preliminary experiments that the results of the methods are similar with that of the existing method using the total station in terms of accuracy. In particular, the method(B) is more efficient than method(A) in that there is no process of estimating location of the conjugate points with the SfM. As a result, through the methods suggested in this study, it was possible to effectively create a 3D model necessary for monitoring the urban regeneration projects, despite of the various constraints that occurs in the low-income neighbourhood environments. But this study was not conducted on all alleyways located in the vicinity of the Sewoon market. This is because there were another environmental constraint which is not treated in this study. For example, in some areas, it was almost impossible to take images with the RTK drone because shading devices attached to façade of buildings completely cover the alleyways or the height of buildings on the side of the alley is too high to take images of alleyways. In these environments, it must be approached in different ways, and leave it as a research project in the future.

ACKNOWLEDGEMENTS

REFERENCES

Taddia, Y., Stecchi, F. and Pellegrinelli, A., 2019. USING DJI PHANTOM 4 RTK DRONE FOR TOPOGRAPHIC MAPPING OF COASTAL AREAS. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.

Tomaštík, J., Mokroš, M., Surový, P., Grznárová, A. and Merganič, J., 2019. UAV RTK/PPK Method—An Optimal Solution for Mapping Inaccessible Forested Areas?. *Remote sensing*, *11*(6), p.721.

Urban, R., Reindl, T. and Brouček, J., Testing of drone DJI Phantom 4 RTK accuracy. In *Advances and Trends in Geodesy, Cartography and Geoinformatics II: Proceedings of the 11th International Scientific and Professional Conference on Geodesy, Cartography and Geoinformatics (GCG 2019), September 10-13, 2019, Demänovská Dolina, Low Tatras, Slovakia* (p. 99). CRC Press.