

## REAL-TIME DRONE MAPPING BASED ON REFERENCE IMAGES FOR VEHICLE FACILITY MONITORING

H. Kim<sup>1,2</sup>, S. Ham<sup>1</sup>, I. Lee<sup>1,\*</sup>

<sup>1</sup> Lab for Sensor and Modeling, Department of Geoinformatics, University of Seoul, 163, Seoulsiripdae-ro, Dongdaemun-gu, Seoul, Korea - (huddy, f100pilot, iplee)@uos.ac.kr

<sup>2</sup> InnoPAM, 77, Cheongpa-ro, Yongsan-gu, Seoul, Korea - huddy1204@gmail.com

Commission II, WG II/1

**KEY WORDS:** Real-time Mapping, Drone, Reference Images, Georeferencing, Aerial Triangulation, Facility Monitoring

### ABSTRACT:

Facilities such as road, parking lots play an important role in our lives nowadays. Damage to such a vehicle facility can cause human injury, as well as inconvenience and cost. To prevent this, facility monitoring is performed periodically, but the current monitoring method is low efficiency by blocking the facility or performing it late at night. In order to increase the efficiency of monitoring, research using images, especially drone images, was conducted. However, when using a drone image, there is a trade-off relationship between accuracy and processing time. In this study, we propose a real-time drone mapping based on reference images for efficient vehicle facility monitoring. The real-time drone mapping based on the reference image is composed of reference images build, aerial triangulation (AT) based on reference images (refAT), and orthophoto generation. The refAT refers to a method of performing AT by using a reference images as reference data. We compared the processing time and processing accuracy of direct georeferencing and refAT. We built 154 drone reference images in the target area. The refAT showed a processing time of about 8.95 seconds and an accuracy of 3.4 cm, and the direct georeferencing method showed a processing time of about 1.49 seconds and an accuracy of 22.5 m. If the method of this study is used for facility monitoring, it is expected that the efficiency of monitoring will be improved with speed and accuracy.

### 1. INTRODUCTION

In modern society, transportation is one of the most essential elements. The state provides public transportation such as buses and subways to ensure the mobility of people who have difficulty using their own transportation. However, even if such a system is in place, the utility is significantly reduced if there are no vehicle facilities such as roads or parking lots. If damage occurs, it has the same effect.

As we can see, accurate and rapid monitoring of facilities with frequent vehicle access such as roads, tunnel and parking lots is important. This is because damage to vehicle facilities not only affects the transportation function of the region, but also affects people's safety. Damage to vehicles such as vehicles inevitably accompanies human injury, and in the event of multiple collisions, the damage increases exponentially. For example, on a snowy day of February 2020, a tank lorry trailer with nitric acid overturned, and about 30 cars following the trailer also collided. As a result, 4 people died and 40 were injured (Park Im-geun & Yerin Choi).

Although it is important to frequently and regularly monitor facilities to prevent damage, it is difficult to perform such monitoring due to the characteristics of vehicle facilities that are difficult to access. Although efforts have been made to proactively prevent damages to facilities through periodic monitoring, the efficiency of transportation has deteriorated since the facilities had to be blocked during the inspection period. In order to minimize the blocking of facilities, monitoring is carried out in the late night or early in the morning, and the effect of monitoring is inevitably low due to low light.

In order to overcome these difficulties, a non-destructive facility monitoring study was conducted to obtain information without directly accessing the facility. Non-destructive inspection is a method that inspects a product or facility without compromising their integrity or surface condition and can save time and money by inspecting facilities without destroying them. Using images is the most representative method to acquire information in non-destructive way.

Usually, various combinations of techniques and sensors such as vibration analysis, 3D reconstruction, vision-based methods were exploited for the purpose of road damage inspection (Kim & Ryu, 2014). Especially, drone photogrammetry became a popular solution for road damage detection due to their ease of use and versatility. Knyaz et al (2016) applied photogrammetric methods to measure road surface and reconstruct the 3D model for road surface management. Díaz-Vilarinho et al (2016) found that photogrammetry using unmanned aerial vehicles (UAVs) could achieve good accuracy and affordable cost for road runoff evaluation. Tan & Li (2019) used UAVs for road condition monitoring, and they could detect road distress from 3D models.

Among them, facility monitoring using drone images is used in many places because it is convenient to acquire data. This is because a drone can monitor a wider range in less time than a method performed by a conventional manpower. Although it is possible to use equipment built in existing infrastructure such as CCTV, drones can actively and immediately check the area desired by the user compared to the fixed surveillance method.

---

\* Corresponding author

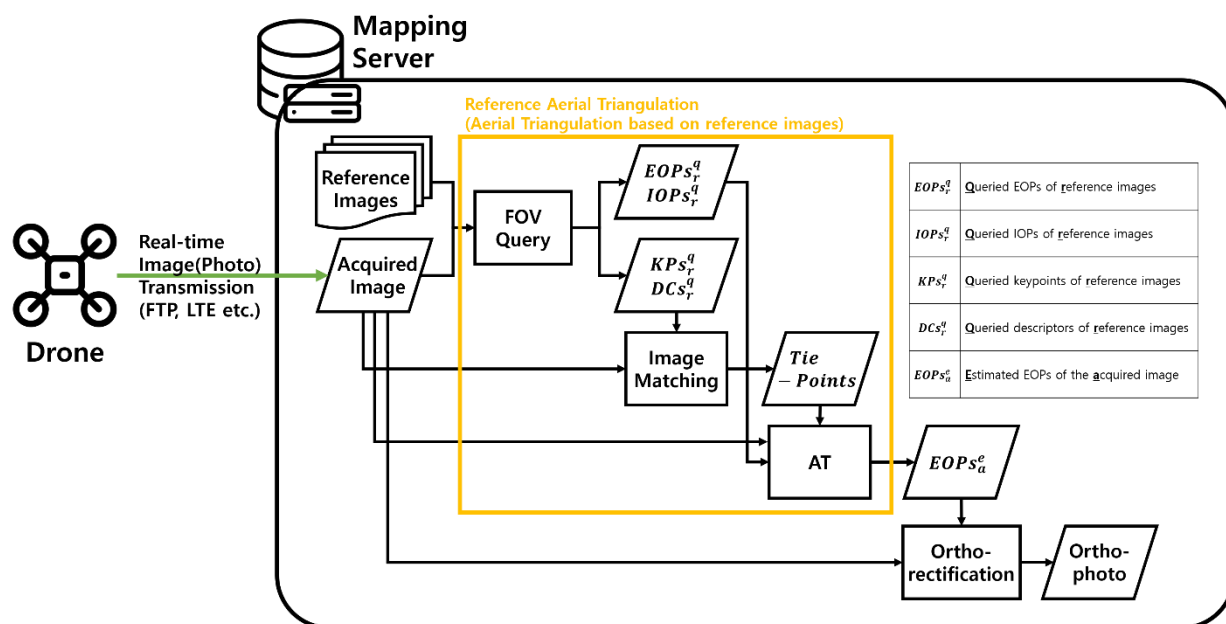


Figure 1. The architecture of proposed method

In general, drone monitoring has been performed based on a post-processing method. Although the post-processing method can guarantee a high level of accuracy, there is an issue that it takes a lot of time for processing. It may yield a large time difference between the final product of the process and the real status of facilities. Therefore, direct georeferencing (DG) that can process data in real-time has been discussed to eliminate this time gap (Bláha et al., 2011). DG is a method of determining the position and orientation of drone images using Global Positioning System (GPS)/Inertial Navigation System (INS) data without adjustment computation. DG has the advantage of short processing time. However, there is a problem that the accuracy of the result is worse.

A method using real-time kinematic (RTK) was proposed to supplement the processing time of the post-processing method and the accuracy of the real-time processing method. RTK uses the information of the reference station to determine the precise location of the rover station in real time, so it can secure cm-level accuracy. An antenna capable of RTK processing can be mounted on the drone to determine the precise position of the drone image and perform accurate monitoring. However, the method using RTK has the difficulty of high cost.

Therefore, in this paper, we propose a real-time drone mapping based on reference images for efficient monitoring of vehicle facilities such as roads and parking lots. Real-time means that we can process data in region of interest (ROI) instantly, mapping based on reference images represents that we can expect the accuracy comparable to post-processing result. We contribute to image-based monitoring on two aspects – processing time and processing accuracy.

## 2. METHODOLOGY

The drone image processing method based on reference images proposed in this study is as follows. First, we build reference images of the target area accurately. It must be done before acquisition of drone images. Then we determine the exterior orientation parameters (EOPs, position/orientation) of the acquired image from a drone through reference aerial

triangulation (AT). Finally, we can generate an individual orthophoto. We are focusing on reference AT, which is determining the estimated EOPs of the image. Figure 1 represent the architecture of our method. Reference AT, the main method, is highlighted.

### 2.1 Aerial Triangulation

AT is one of georeferencing methods that determine the position and orientation of a camera at the moment when an image is acquired for two or more images. Among them, bundle adjustment based on the collinearity equation is the most widely used. Bundle adjustment is a method of collectively adjusting the bundle using the image points projected from the same ground point of the acquired images, and thereby determining the EOPs and the coordinates of ground points of the image. Bundle means a set of rays connecting the image point, the projection center of the image, and the ground point. The collinearity equation indicates that the image point, the projection center of the image, and the ground point exist on one straight line (Mikhail, 2001). The bundle adjustment sets the error between the computed coordinate of the image point projected at the ground point and the observed coordinates of the image point as the objective function, and adjusts the EOPs and the ground point of the image so that the objective function is minimized. If the internal orientation parameters (IOPs) of the camera is a constant, the mathematical model of bundle adjustment can be expressed as equation (1).

$$(x_i^j, y_i^j) = f(EOP_1, \dots, EOP_m, X_1, Y_1, Z_1, \dots, X_n, Y_n, Z_n) \quad (1)$$

where  $m$  = the number of images  
 $n$  = the number of ground points  
 $(x_i^j, y_i^j)$  = coordinates of image points

### 2.2 Structure of reference images

Reference images refer to images in which elements necessary for performing AT are precisely determined. In order to determine the EOPs of the image through the AT, tie points and reference data between the images are required. The amount and quality of the tie points and reference data is important because

it greatly affects the accuracy of the EOPs of the resulting AT image. The tie points consist of keypoints, which are feature points of the image, and descriptors, which are information about it. Reference data consists of control points and EOPs (including IOPs). That is, the reference images are composed of (1) keypoints, (2) its descriptor, (3) EOPs, and (4) IOPs of accurately processed images.

Each of the elements of the reference images is used in the next step. First, EOPs are used in a field of view (FOV) query to find reference images corresponding to processing range of an acquired image, and then keypoints and descriptors are used in image matching with an acquired image. In AT, EOPs and IOPs are used to adjust. Table 1 shows which element is used in which processing steps.

	FOV Query	Image Matching	Aerial Triangulation
Keypoints	-	O	O (Tie points)
Descriptors	-	O	
EOPs	O	-	O
IOPs	-	-	O

Table 1. Processing steps and each element

### 2.3 Reference Aerial Triangulation

The refAT proposed in this study differs from the existing AT in (1) tie points extraction and (2) EOPs adjustment. First, it is possible to improve the accuracy in the tie points extraction. Since a candidate reference image to be processed and an acquired image are selected before image matching, mismatches that may occur due to similarity of features in images can be minimized even in images that are not related to each other in position. Further, the processing time can be shortened. In general image matching, keypoints are extracted from all images, but in this method, keypoints of the reference images are extracted in advance and only keypoints of the acquired image are extracted.

Next, the EOPs are used as a constraint in the adjustment, not control points. Equation (1) can be expressed as equation (2) with a Gauss-Markov Model. If this is expressed by dividing the EOPs part  $A_e$  and the ground point coordinate part  $A_g$ , it is as shown in equation (3).

$$Y - f(\Xi_{e0}, \Xi_{g0}) = \left[ \frac{\partial f}{\partial \Xi_e} \quad \frac{\partial f}{\partial \Xi_g} \right] \begin{bmatrix} \xi_e \\ \xi_g \end{bmatrix} + e \quad (2)$$

where  $Y$  = image coordinates  
 $\Xi_e$  = EOPs  
 $\Xi_g$  = ground coordinates  
 $\Xi_{e0}$  = initial value of EOPs  
 $\Xi_{g0}$  = initial value of ground coordinates  
 $\xi_e, \xi_g$  = difference between value of initial and true  
 $e$  = measurement error of image coordinates

$$y = [A_e \quad A_g] \begin{bmatrix} \xi_e \\ \xi_g \end{bmatrix} + e \quad (3)$$

Since equation (3) does not contain information about ground coordinates, it has seven rank deficiencies. To compensate for this, this study sets the EOPs of the reference images to the constraints. This is shown in equation (4).

$$\begin{bmatrix} y_i \\ y_e \end{bmatrix} = \begin{bmatrix} A_e & A_g \\ K_e & 0 \end{bmatrix} \begin{bmatrix} \xi_e \\ \xi_g \end{bmatrix} + \begin{bmatrix} e \\ e_0 \end{bmatrix}, \begin{bmatrix} e \\ e_0 \end{bmatrix} \sim \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} P_i^{-1} & 0 \\ 0 & P_e^{-1} \end{bmatrix} \right) \quad (4)$$

where  $y_i$  = difference between measured and computed value of image coordinates  
 $y_e$  = difference between measured and computed value of EOPs and constraints  
 $A_e$  = design matrix for EOPs  
 $A_g$  = design matrix for ground coordinates  
 $K_e$  = design matrix for constraints of EOPs  
 $e$  = measurement error of image coordinates

$A_e$  and  $A_g$  are established by differentiating the collinear equation for each variable ( $A_e$ : EOPs,  $A_g$ : ground coordinates), and  $K_e$  is set as an identity matrix because the external expression elements of the reference image are used as direct constraints. Since the EOPs of the acquired image is not used as a constraint,  $K_e$  is established only for the reference images, not the acquired image. Therefore, the length of the row of  $K_e$  is proportional to the number of reference images, and is not affected by the number of acquired images.

## 3. EXPERIMENTAL RESULTS

Chapter 3 presents experimental results to verify the methodology of this study. We describe the datasets, including the experimental data and the acquisition area, and verify the results of processing the datasets. The processing result is divided into processing time and processing accuracy. 1) In the verification of processing time, the method of this study is compared with the DG method. 2) In the verification of processing accuracy, the accuracy of the method of this study and the real-time processing method are compared.

### 3.1 Datasets

We acquired drone image dataset A and B in January 2018 and April 2018 respectively. The images of both datasets were taken using DJI Mavic PRO drone, and they were acquired in the typical rural area of Yangpyeong-gun, Gyeonggi-do, Korea. Figure 2 shows the target area in detail.

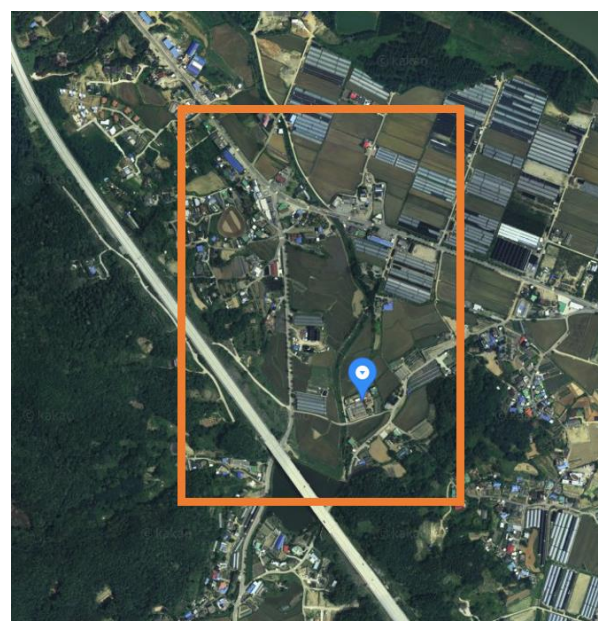


Figure 2. Test area



Dataset A consists of 10 images, and dataset B consists of 154 images. Especially, in the images of dataset A, some snow-covered area can be observed since the photos were taken in the middle of winter. Dataset A is used as 'acquired image' and dataset B is used as 'reference images'. The altitude was 120 m from the home point, and the ground sampling distance was 3.9 cm. Figure 3 and figure 4 shows some example of the acquired images. Also, we surveyed 8 ground control points to perform aerial triangulation and positional accuracy analysis.



Figure 3. Examples of dataset A (acquired images)



Figure 4. Examples of dataset B (reference images)

Using the image dataset B, we performed structure from motion (SfM) and multi-view geometry (MVG) using Agisoft Metashape. The root means square error (RMSE) of ground control points was 9 cm. The generated dense point cloud and positions of the ground control points are shown in figure 5.



Figure 5. The generated point cloud and distribution of GCPs

### 3.2 Time verification

We measured processing time of 10 images of dataset A. Not only the processing time of refAT was measured, but also the processing time of DG was measured as well to compare performance. As it is described in table 2, the processing time of refAT was longer than DG, but it was shorter than typical photogrammetric process which requires few minutes to hours.

# of Image	refAT (sec)	DG (sec)
1	9.29	2.21
2	9.06	1.68
3	8.97	1.39
4	9.05	1.35
5	8.85	1.37
6	9.02	1.35
7	8.64	1.32
8	8.59	1.41
9	9.06	1.44
10	8.97	1.33
Mean	8.95	1.49
Std	0.21	0.27

Table 2. Processing time of refAT and DG

### 3.3 Accuracy verification

We analyzed positional accuracy of our method. Our goal of accuracy verification is to check our method (refAT) performs better than DG. The positional accuracy was measured through comparing image points and corresponding ground control points that we had already surveyed. We calculated mean, standard deviation and RMSE about x-axis and y-axis. As described in table 3 and table 4, the RMSE of refAT was better than DG. DG was good in terms of processing time, but the positional accuracy was not good to satisfy our research purpose. On the other hand, refAT was bit slower than DG but the positional accuracy was satisfying to be exploited as vehicle facility monitoring.

Figure 6 and figure 7 depict individual orthoimages generated by DG and refAT respectively. The locations of the check points are overlaid as triangular points, and the image points corresponding to the check points are overlaid as circular points. Colors of the check points and image points are matched corresponding point IDs. The image points of refAT are closer to the check points, while that of DG are more scattered.



Figure 6. Check points and corresponding image points (DG)



Figure 7. Check points and corresponding image points (refAT)

Table 3 and table 4 describe the positional accuracy of DG and refAT respectively. The RMSE of refAT was smaller than DG. Also, since the standard deviation of refAT is lower than DG, each individual ortho-images of refAT are relatively well-registered.

$\mu_x$ ( $\sigma_x$ , m)	$\mu_y$ ( $\sigma_y$ , m)	RMSE (x, m)	RMSE (y, m)	RMSE (XY, m)
0.745 (3.978)	-0.460 (3.994)	16.018	15.799	22.499

Table 3. Positional accuracy of DG

$\mu_x$ ( $\sigma_x$ , m)	$\mu_y$ ( $\sigma_y$ , m)	RMSE (x, m)	RMSE (y, m)	RMSE (XY, m)
0.118 (0.081)	-0.112 (0.120)	0.020	0.027	0.034

Table 4. Positional accuracy of refAT

The mean, standard deviation, and RMSE of refAT all showed better results than DG. First, the average of error of ground points in DG was about 75 cm and -46 cm in the order of x and y axes, and that in refAT was about 11 cm. The bias for the experimental value showed a larger value in DG. In the standard deviation, DG showed a value close to 4 m, so the relative accuracy was also lower than that of refAT, which is about 10 cm. The bias and relative accuracy of the result value affects the RMSE, and it is considered that the following result was shown.

We also analyzed the results of our method in qualitative way. In figure 8, which is a part of DG individual orthophotos, we could find seamlines that it makes harder to identify facilities. However, in figure 9, which is the same part of refAT individual orthophotos, the seamlines look better than DG. It makes easier to identify facilities.



Figure 8. Seamlines between individual orthoimages (DG)



Figure 9. Seamlines between individual orthoimages (refAT)

#### 4. CONCOLUSIONS

In this paper, we suggest a real-time drone mapping method for rapid and efficient vehicle facility monitoring. There have been many attempts to use drones for vehicle facility monitoring, but the existing drone mapping has a trade-off relationship between processing time and accuracy. To compensate for this, we proposed a method to improve the drone mapping processing time and accuracy using a reference image.

Reference images are datasets that precisely determines the elements necessary for positioning the image in advance and has the following elements. Definition of relative relationship between images: (1) keypoints and (2) descriptor, reference data when adjusting: (3) EOPs and (4) IOPs. When adjusting, the constraint design matrix was designed to use only EOPs of the reference image.

The method of this study was verified in terms of processing time and accuracy. In Yangpyeong-gun, Gyeonggi-do, drone images were acquired twice in spring and winter. 154 images taken in spring were built as reference images, and 10 of winter images were used as query images. First, the processing time was about 8 seconds per image, which showed a difference of about 6



seconds per image when compared to the fastest processing method, DG. In terms of accuracy, refAT has an accuracy of about 3.4 cm and DG has an accuracy of 22.5 m.

The method of this study maps the drone image to cm-level accuracy within seconds. Compared with the post-processing method, which takes tens of minutes or more, and DG, which shows m-level accuracy, it shows improved results. However, in order to approach the accuracy of the post-processing method and the speed of the DG, the following items need to be improved. First, in order to improve the processing time, detailed classification and calculation of the reference image participating in the processing with the acquired image are necessary. In addition, the use of DEMs in the region will make progress in terms of processing time and accuracy. It is expected that the method of this study can be used in various environments that require fast and efficient monitoring.

## REFERENCES

Park Im-geun & Yerin Choi. (17.02.2020). Multi-collision fire accident at Suncheon ~ Wanju Expressway Tunnel... 4 people died, and 40 others were injured. Hankyoreh. Retrieved from <http://www.hani.co.kr/arti/area/honam/928536.html>.

Bláha, M., Eisenbeiss, H., Grimm, D., & Limpach, P. (2011). Direct georeferencing of UAVs. In Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g) (Vol. 38, pp. 131-136). Copernicus.

Díaz-Vilariño, L., González-Jorge, H., Martínez-Sánchez, J., Bueno, M., & Arias, P. (2016). Determining the limits of unmanned aerial photogrammetry for the evaluation of road runoff. *Measurement*, 85, 132-141.

Kim, T., & Ryu, S. K. (2014). Review and analysis of pothole detection methods. *Journal of Emerging Trends in Computing and Information Sciences*, 5(8), 603-608.

Knyaz, V. A., & Chibunichev, A. G. (2016). Photogrammetric techniques for road surface analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 41(B5), 515-520.

Tan, Y., & Li, Y. (2019). UAV Photogrammetry-Based 3D Road Distress Detection. *ISPRS International Journal of Geo-Information*, 8(9), 409.

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

Mikhail, E. M., Bethel, J. S., McGlone, J. C. (2001). Introduction to modern photogrammetry (Vol. 1), John Wiley & Sons Inc, pp. 119-123.