GEOMETRIC CORRECTION OF SMARTPHONE IMAGES USING UAVs REFERENCE IMAGES FOR FACILITY MAINTENANCE

Hwiyoung Kim, Kyoungah Choi, Impyeong Lee
Lab. for Sensor & Modeling, Department of Geoinformatics, University of Seoul,
Seoulsiripdaero 163, Dongdaemun-gu, Seoul 02504, Korea,
Email: huddy@uos.ac.kr, shale@uos.ac.kr, iplee@uos.ac.kr

KEY WORDS: geometric correction, smartphone images, UAVs, reference images, image-based maintenance

ABSTRACT: Continuous facility maintenance is important to prevent accidents. Currently, the method of facility maintenance is dependent on visual inspection, which is labor-intensive, time-consuming and has a risk of casualties. In addition, the results of facility inspection can be different according to inspectors. In this paper, we propose an image-based inspection method, which measures location of points of interest using a smartphone image for effective facility maintenance. We performed georeferencing of smartphone images using pre-processed unmanned aerial vehicles (UAVs) reference images, not ground control points (GCPs). We then rectified the images using position of smartphone images derived from georeferencing onto a target plane of the facility. Experimental results show that the proposed approach has accuracies of 3.543 cm, 0.209° in the position of smartphone images and 9.829 cm in the location of points measured from images. We expect that this method can enhance utilization of image based facility maintenance.

1. INTRODUCTION

Facilities such as roads and bridges play important role in our lives. If such a facility is damaged or becomes unavailable, it will cause a great inconvenience. Moreover, if the facility collapsed, there will be an enormous cost to rebuild the damage as well as the loss of life. The collapse of the Pfeiffer Canyon Bridge in Big Sur, California, in March 2017, took six hours to travel between adjacent areas and resulted in an estimated recovery cost of approximately $ 20M(Brinklow, A., 2017). In October 2016, the Lecco overpass collapsed in the Province of Lecco. Overpass covered the road and the vehicle that was passing the road was totally destroyed, resulting in one death and five injured(BBC NEWS, 2016). In order to prevent accidents of facilities, maintenance should be continuously carried out.

Currently, the method of facility inspection depends on manpower. The inspector approaches the facility and visually checks the condition of the facility. However, visual inspection is labor-intensive and time-consuming because the inspectors approach the facility one by one, and the risk of the inspectors is high. In addition, since the condition of the facility is visually judged, the inspection result may be changed according to the inspector's subjectivity. This hinders efficient and objective inspection of facilities.

In order to solve this problem, researches have been made on the use of equipment such as sensors. Particularly, the use of a vision sensor such as a camera has been studied to acquire object information. A study was conducted to classify bolt, anchor, spike, etc(Resendiz *et al.*, 2013; Feng *et al.*, 2014). Research has also been carried out to detect damage such as cracks through processes such as binarization and conversion of captured images of bridges(Oh *et al.*, 2009; Adhikari *et al.*, 2014; Yeum *et al.*, 2015). Also, the tension of the cable was obtained by using the image of the cable-stayed bridge(Kim *et al.*, 2013). 3-dimensional information of facilities was reconstructed(Lattanzi and Miller, 2014; Jahanshahi and Masri, 2012) and used for facility inspection. 3D information was reconstructed using stereo image of facility, 3D camera and laser scanner(Jarzebowicz and Judek, 2014)

One of the important parts of the image-based method is the efficient acquisition of images. Even if the image-based method itself has efficiency, if the acquisition of the image is difficult, the efficiency will be reduced. Most people are using smartphones recently. According to Google's 'Mobile Apps in APAC: 2016 Report', smartphone penetration is higher than desktop penetration. In Korea, the penetration rate is 80% in 2015 and 90% in 2016 (Google, 2016). Along with high penetration rates, smartphones have the advantage of being portable. That is, if the smartphone camera is used as the image acquiring means, the image can be acquired at a desired time and place. This makes it possible to efficiently acquire images and to improve the efficiency of the image-based method.

In this study, we propose a smartphone image-based method for effective facility maintenance. Image-based maintenance consists of georeferencing and rectification steps. Georeferencing used reference images for acceleration and automation instead of the commonly used ground control points(GCPs). To perform rectification, two or more images taken at the same point should be used. In this study, a single image was rectified by defining the projection plane of the image. In Chapter 2, the whole flow of the image-based inspection method and the explanation of georeferencing and rectification are presented. Section 3 shows details of the data acquisition area and specifications, and Section 4 shows the results of the processing. Chapter 5 summarizes the research and discusses limitations and future works.

2. METHODOLOGY

This chapter presents a methodology for using images to maintain facilities. The methodology consists of georeferencing and rectification steps. Exterior Orientation Parameters(EOPs) of the image are acquired in the process of georeferencing are determined based on reference images, and next the image is projected and resampled in the region of interest of the facility using the determined EOPs. Figure 1 shows the flow of the facility maintenance method using images and each component.

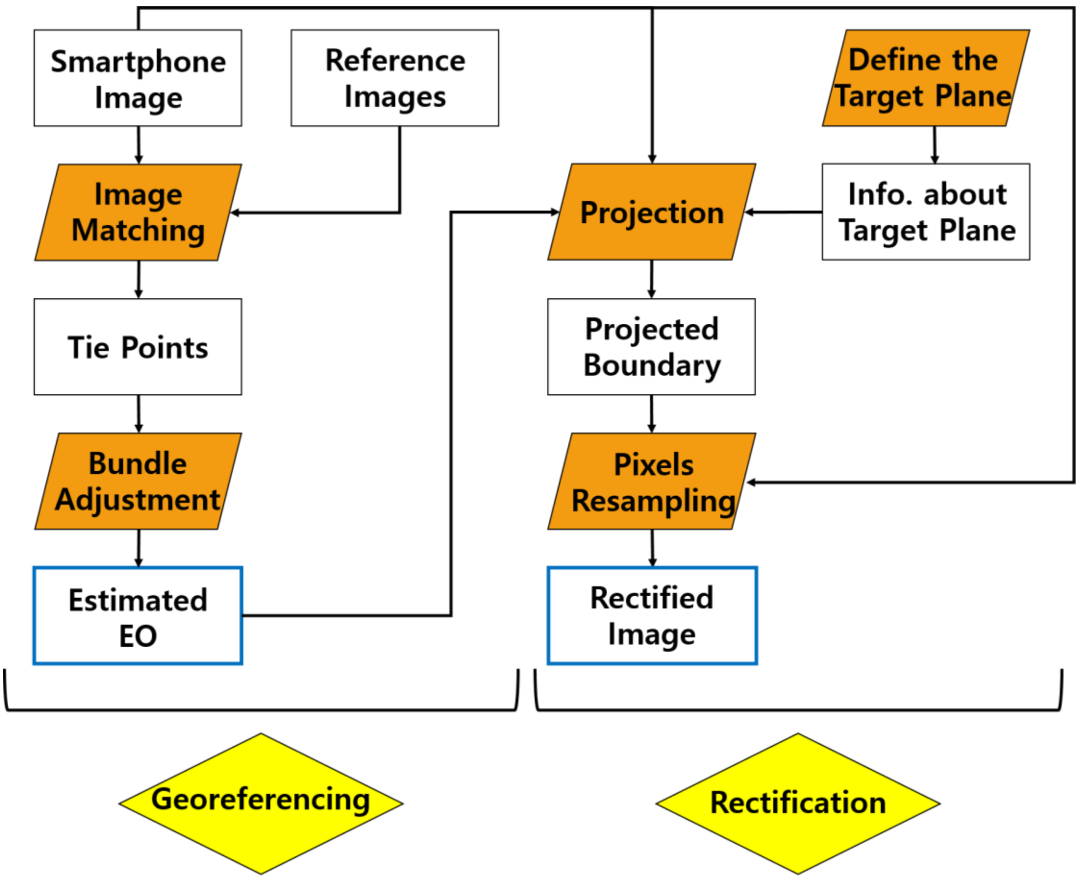

Figure 1. Flowchart of proposed method for georeferencing and rectification of smartphone images based on reference images

2.1 Georeferencing

Georeferencing is the process of determining EOPs when acquiring images. Georeferencing is divided into image matching, which calculates the conjugation point between images, and bundle adjustment, which determines the precise EOPs of the image using the initial EOPs of the image and the conjugation point calculated through image matching. The image matching consists of ‘detector’ that detects the feature points of the image and ‘descriptor’ that shows the information of the detected feature points. The feature points detected in each image through the detector have a descriptor which is a feature vector representing their information, and the feature vectors of the feature points are compared to connect the points determined to represent the same information. The bundle adjustment is based on collinearity equation that the perspective center of the image point, the image point where the ground point is projected and the ground point are on a straight line. In bundle adjustment, a bundle refers to a bundle of collinear lines. The bundle is a set of collinear lines. The bundle is used to obtain the EOPs of the image that minimizes the difference between the position of the corresponding image point observed in the image and the position of the image point projected from the ground point.

In order to perform bundle adjustment, initial EOPs of images, tie points between images, and reference data are required. The initial EOPs of the image are needed because mathematical model of the collinearity equation is nonlinear, and tie points are necessary to establish the relationship between images. The reference data is necessary to represent the image in the absolute coordinate system of the ground. Generally, the reference data is obtained by using the GCPs obtained by using the global positioning system(GPS) receiver or total station. However, GCPs require much time and cost to acquire. However, since GCPs require much time and cost to acquire, it may be difficult to efficiently perform georeferencing even if the image is acquired quickly and easily using a smartphone camera. Using reference images instead of GCPs as the reference data can compensate for these drawbacks. Reference images means an image in which the EOPs of the image and Interior Orientation Parameters (IOPs) are precisely determined in advance. If reference images are used as the reference data, it is not necessary to acquire GCPs every time the bundle adjustment is performed, and the process of assigning GCPs to the image can be omitted, so that the EOPs of the image can be determined only by the acquired image. The input values of the method using GCPs as the reference data and the method using reference images are shown in Figure 2.

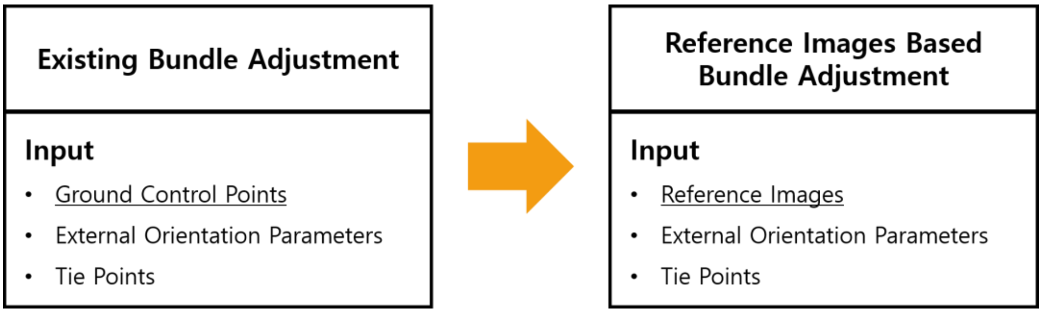

Figure 2. Comparison of existing bundle adjustment and reference image based bundle adjustment

However, the bundle adjustment using reference images must satisfy the minimum condition. In order to establish the relative relationship between adjacent images, at least five tie points are required. In order to refer to the image as an absolute coordinate system, seven parameters defining the coordinate system should be determined. A relative relationship can be established through tie points between the smartphone image and the reference image, but the scale in the absolute coordinate system cannot be determined. To determine this, the coordinates of the ground points are required, and the coordinates of the ground points can be confirmed by the tie points identified in two or more reference images. Therefore, in order to perform bundle adjustment based on reference images proposed in this study, at least five tie points are required between one smartphone image and the reference image, and at least one of the tie points should be confirmed on two or more reference images. As shown in Figure 3, the bundle adjustment can be performed based on reference images if the smartphone image has a reference image and at least five tie points, and one of the points has a tie point with another reference image.

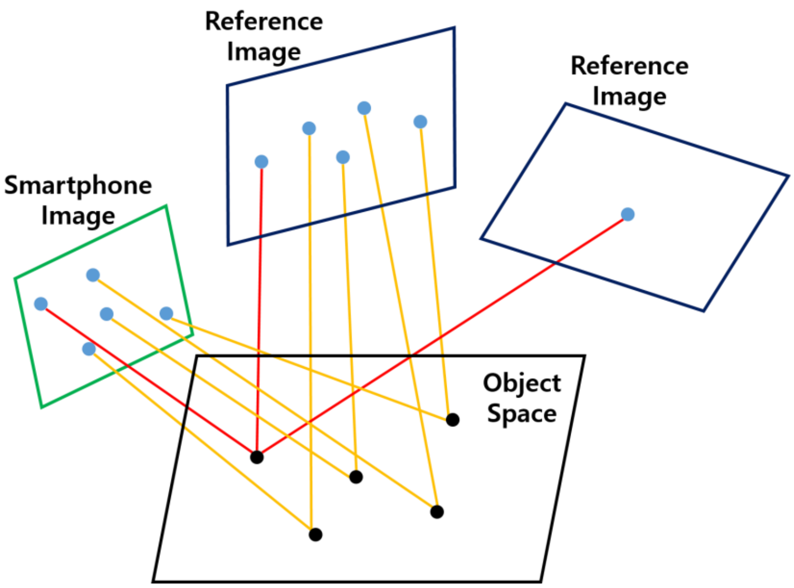

Figure 3. The minimum condition of reference images based bundle adjustment

2.2 Rectification

Images are acquired by the center projection. The process of transforming this into an orthographic projection and referring to the absolute coordinate system is referred to as rectification. In this study, a smartphone image is projected on the target plane of the facility to perform rectification. The concept of this is shown in Figure 4. Projection of the image is based on a collinearity equation as bundle adjustment. The smartphone image is projected on the object plane using the collinearity equation to calculate the coordinates of the ground point on the object plane, and the pixels of the smartphone image are resampled.

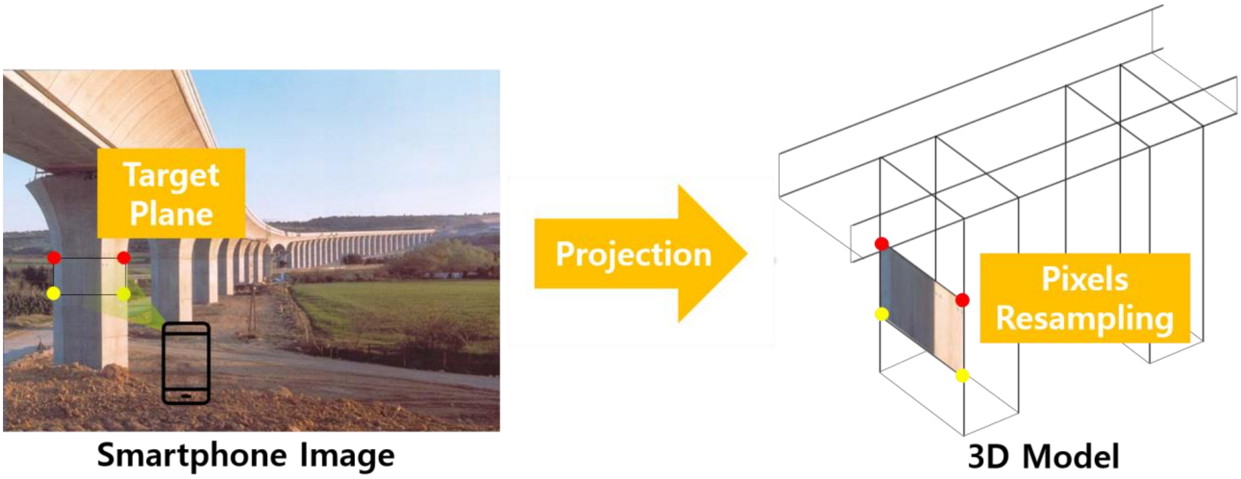

Figure 4. The concept of rectification onto target plane

To determine the coordinates of a ground point using an image, the ground point must be observable from two or more images. Since the distance from the image to the ground cannot be known with only one image, the coordinates of the ground point on which the image is projected cannot be determined. In this study, we solve this problem by setting the object plane on which the image is projected. The distance from the image to the object plane is calculated, and the coordinates of the ground point on which the image is projected are calculated. The method of obtaining the distance from the image to the ground is as follows.

When the collinearity equation is expressed as Equation (1), λ represents the scale of the image and the distance from the image to the ground. If we define the object plane to project the smartphone image as Equation (2), we rearrange Eq. (1) for arbitrary ground points P_G shown in the smartphone image and replace the summed P_G with Equation (2) Equation (3) can be rearranged by using the EOPs of the image and the information of the object plane. Using the calculated λ, the ground point coordinates on the object plane on which the image points of the smartphone image are projected can be calculated as Equation (4).

$$P_C = \frac{1}{\lambda} R_{GC}(P_G - C_G) \tag{1}$$

where P_C: the image point in Camera Coordinate Systems (CCS)
λ: the ratio of length between image and ground
R_{GC}: the rotation matrix converting Ground Coordinate Systems (GCS) to CCS
P_G: the ground point in GCS
C_G: the perspective center in GCS

$$n^T p = d \tag{2}$$

where n: the normal vector of a plane
p: a point on the plane
d: the value of a point calculated by plane equations

We can derive the scale using Equation (3) as p in Equation (2) corresponds to P_G in Equation (1). Equation (3) represents the process of deriving the scale. As shown in Equation (4), ground coordinates of the projected image can be obtained.

$$\begin{gathered} n^T(\lambda R_{CG}P_C + C_G) = d \\ \lambda n^T R_{CG}P_C = d - n^T C_G \\ \lambda = \frac{d - n^T C_G}{n^T R_{CG} P_C} \end{gathered} \tag{3}$$

where R_{CG}: the rotation matrix converting CCS to GCS

$$\begin{gathered} P_G = \lambda R_{CG}P_C + C_G \\ P_G = \frac{d - n^T C_G}{n^T R_{CG} P_C} R_{CG}P_C + C_G \end{gathered} \tag{4}$$

The coordinates when the image is projected onto the ground can be determined using Equation (4). If the coordinates of the ground points are determined for all the pixels of the image, rectification can be performed. However, since it requires an excessive amount of calculation, it is necessary to interpolate the coordinates of pixels projected at intervals of resolution by designating a range in which the image is projected. After setting the range in which the smartphone image is projected on the object plane, a grid is generated at intervals of the resolution in the corresponding range. Next, we resample the pixels by back-projecting the ground point coordinates of each grid to the image. However, since the object plane has arbitrary directionality with respect to the three axes defining the ground coordinate system, it is not useful to interpolate the ground point coordinates in each grid. In this paper, we define the object plane coordinate system, calculate the ground point coordinates for the lattice based on the object plane coordinate system, perform the image resampling, and convert it to the ground coordinate system. The coordinates of an arbitrary point $\vec{a}$ in a coordinate space with a basis of $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\}$ can be defined as Equation (5). Using the normal vector of the object plane as a base, the ground coordinates of the grid can be expressed by Equation (6) based on the object plane coordinate system. By multiplying the base of the object plane coordinate system by Equation (6), we can convert it back to the ground coordinate system.

$$\vec{a} = \mathrm{p}\overrightarrow{v_1} + q\overrightarrow{v_2} + r\overrightarrow{v_3} = [\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad \overrightarrow{v_3}][p \quad q \quad r]^T \tag{5}$$

where $\vec{a}$: an arbitrary point
$\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$: the basis of an coordinate system
p, q, r: the coordinates of $\vec{a}$ in the basis of $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\}$

$$[\overrightarrow{n_1} \quad \overrightarrow{n_2} \quad \overrightarrow{n_3}]^{-1}\vec{a} \tag{6}$$

where $\overrightarrow{n_1}, \overrightarrow{n_2}, \overrightarrow{n_3}$: the basis of surface coordinate system
$\vec{a}$: an arbitrary point

3. DATA CONSTRUCTION

3.1 Data Acquisition

The data used in this study were obtained from Gangjeong-goryeong-bo in Daegu, South Korea. As shown in Figure 5 (a), smartphone images were taken in the range of about 70 m from the entrance of Gangjeong-goryeong-bo. It is difficult for people to approach it due to its characteristics. Therefore, in order to take the side of the bridge it must be taken looking up from below, as in Figure 5 (b). Table 1 shows the specifications of the smartphone cameras. The number of smartphone images that captured the bridge is 12. 7 images were from TG-L800S, 5 images were from iPhone 7.

(a) (b)

Figure 5. (a) Acquisition range of smartphone images, (b) a smartphone image

Table 1. Specification of smartphone cameras

Model	Focal length	Image resolution	Pixel size
iPhone 7	3.99 mm	4032 × 3024 px	1.22 um
TG-L800S	3.85 mm	4160 × 2340 px	1.25 um

3.2 Reference Images

To obtain accurate results in image analysis, high-resolution images should be used. The resolution of the image increases as the camera approaches the object. However, facilities such as Gangjeong-goryeong-bo are difficult to access and there is a limitation in acquiring high resolution images. There have been attempts to acquire images of facilities using UAVs with high mobility. (Tokmakidis and Scarlatos, 2002, Kumar et al., 2013, Delair-Tech SAS, 2014). In this study, a camera was installed in UAVs to obtain high accuracy of reference images used as reference data.

The acquisition range of reference images is similar to that of smartphone images. However, since reference images were acquired at the same height as the Gangjeong-goryung-bo using UAVs, the angle of view is different from that of smartphone images taken from above the Gangjeong-goryung-bo. Table 2 shows the specifications of the cameras used for reference images acquisition, and UAVs images are shown in Figure 6. GCPs were used to determine the precise EOPs and IOPs of reference images. 301 images were processed using 32 GCPs and 6 check points(CPs). About 4 or 5 GCPs are distributed per image, and the accuracy of CPs is about 1.48 cm.

Figure 6. UAVs reference images

Table 2. Specification of a camera used for acquiring reference images

Model	Focal length	Image resolution	Pixel size
ILCE-7RM2	35 mm	7952 × 5304 px	4.53 um

4. EXPERIMENTAL RESULTS

4.1 Georeferencing

Generally, the accuracy of CPs is used to evaluate the accuracy of georeferencing. However, since georeferencing based on reference images does not use ground points such as the CP, the accuracy is evaluated in a different way. EOPs of the images obtained by performing georeferencing using GCPs are set as reference values and compared with the EOPs of the images obtained by performing georeferencing based on reference images. The number of GCPs and CPs is the same as that of reference images.

It is important to consider the IOPs of the image when performing georeferencing. Correction of distortions in the image can lead to higher accuracy. In this study, distortion of reference image and smartphone image should be corrected. IOPs of the reference image were estimated by self-calibration when constructing the reference image. Self-calibration is a method of obtaining EOPs of images and IOPs of images together when performing georeferencing. Due to the nature of smartphone cameras, it is difficult to estimate IOPs accurately due to frequent changes of IOPs while taking photos. Therefore, IOPs of smartphone images are assumed to be free of distortion when performing georeferencing.

When the georeferencing was performed using GCPs, the accuracy of CPs was 1.46 cm, which is similar to the accuracy of CPs of reference images. We calculated the RMSE with EOPs calculated by performing georeferencing based on reference images by setting this value as reference values. As shown in Table 3, it represents 3.543 cm for the position, and 0.209 ° for the orientation.

Table 3. EOPs RMSE of smartphone images

X (cm)	Y (cm)	Z (cm)	Total (cm)	ω (°)	ϕ (°)	κ (°)	Total (°)
1.852	1.564	2.585	3.543	0.168	0.088	0.089	0.209

4.2 Rectification

In order to rectify a single image, EOPs of each image and information of the target object plane to be projected are required. EOPs are acquired through georeferencing and the information of the target object plane is obtained by defining the equation of the plane on which the image is projected. The plane equation can be represented by at least three points on a plane that are not on the same straight line. As shown in Figure 7, the plane equation is defined using three points (P1, P2, P3) in the image. Two vectors $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ on the plane can be set using three points, and the normal vector $\vec{n}$ of the object plane can be defined by cross product of these two vectors.

Smartphone images were projected on the target object plane by substituting the defined $\vec{n}$ and EOPs of each image into Eq. (4). The projected boundary of each smartphone image was calculated by projecting four vertices of the smartphone image, and the projected coordinates of the image were calculated and interpolated. When interpolating the projected coordinates, we used the normal vector $\vec{n}$ of the object plane defined when projecting the image. Then, pixels of the image were resampled as when the image was projected. Figure 8 shows an example of rectified images. We verified the accuracy of rectification comparing the survey points and computed points from rectified images. We computed RMSE of 7 points. The accuracy represented 9.829 cm, as shown in Table 4.

Figure 7. Information of a target object plane

Figure 8. Rectified images

Table 4. RMSE of ground points from rectified images

X error (cm)	Y error (cm)	Z error (cm)	Total (cm)
7.366	5.949	2.635	9.829

5. CONCLUSIONS

In this study, we propose a method to use a smartphone image to inspect the facilities efficiently. At present, the inspector directly accesses the facility and visually inspects the condition. The image-based method has an efficiency of inspection in terms of cost, time, and safety as compared with the existing method. In particular, since smartphones have a high penetration rate and are portable, acquiring images using smartphone cameras can improve the efficiency of facility inspection. When georeferencing the acquired smartphone image, the time and cost required to use GCPs by using reference images could be saved as the reference data. In addition, reference images were constructed using UAVs considering characteristics of experimental area where human access is difficult. The EOPs of the smartphone images obtained through georeferencing showed that the images were rectified to the target plane of the facility, and the location of the facility was indicated by the images. Experimental results show that 12 smartphone images are processed and the accuracy of EOPs is 3.543 cm, 0.209 deg. The accuracy of the estimated ground points was 9.829 cm.

What is important in the inspection of the facilities is to determine what kind of damages occurred to the facilities and to what extent. Therefore, in order to maximize the utilization of the image in the facility inspection, it is necessary to grasp the type of damage occurred in the facility such as cracks or scrapes through the acquired image, and to quantify the damage such as the length of the crack or the size of the whitewash. In this study, only the accuracy of quantification was evaluated, but it is also necessary to provide information on the type of damage in the future.

ACKNOWLEDGMENTS

This work was supported by the Building Wide Integrated Surveillance System of Marine Territory project (1525004602) funded by Republic of Korea government and Ministry of Oceans and Fisheries.

REFERENCES

Adhikari, R. S., Moselhi, O., & Bagchi, A. 2014. Image-based retrieval of concrete crack properties for bridge inspection. Automation in construction, 39, 180-194.

BBC NEWS, 2016. Italian bridge collapses on busy road in Lecco, Retrieved September 23, 2017, from http://www.bbc.com/news/world-europe-37810808#

Brinklow, A., 2017. Big Sur bridge damaged by weather comes crashing down, Retrieved September 23, 2017, from https://sf.curbed.com/2017/3/20/14983602/pfeiffer-canyon-bridge-collapse

Delair-Tech SAS, 2014. Power lines inspection using mini-uav, Delair-Tech, Madrid, Spain.

Feng, H., Jiang, Z., Xie, F., Yang, P., Shi, J., & Chen, L. 2014. Automatic fastener classification and defect detection in vision-based railway inspection systems. IEEE transactions on instrumentation and measurement, 63(4), 877-888.

Google, 2016, Mobile Apps in APAC: 2016 Report, Retrieved September 23, 2017, from http://apac.thinkwithgoogle.com/intl/en/articles/mobile-apps-in-apac-2016-report.html

Jahanshahi, M. R., & Masri, S. F. 2012. Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. Automation in Construction, 22, 567-576.

Jarzebowicz, L., & Judek, S. 2014, 3D machine vision system for inspection of contact strips in railway vehicle current collectors. In Applied Electronics (AE), 2014 International Conference on (pp. 139-144). IEEE.

Kim, S. W., Jeon, B. G., Kim, N. S., & Park, J. C. 2013. Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge. Structural Health Monitoring, 12(5-6), 440-456.

Kumar, K., Rasheed, A., Kumar, R., Giridharan, M., Ganesh, M., 2013. Dhaksha, the unmanned aircraft system in its new avatar-automated aerial inspection of India’s tallest tower, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W2.

Lattanzi, D., & Miller, G. R. 2014. 3D scene reconstruction for robotic bridge inspection. Journal of Infrastructure Systems, 21(2), 04014041.

Oh, J. K., Jang, G., Oh, S., Lee, J. H., Yi, B. J., Moon, Y. S., Lee, J. S., & Choi, Y. 2009. Bridge inspection robot system with machine vision. Automation in Construction, 18(7), 929-941.

Resendiz, E., Hart, J. M., & Ahuja, N. 2013. Automated visual inspection of railroad tracks. IEEE Transactions on Intelligent Transportation Systems, 14(2), 751-760.

Tokmakidis, K., Scarlatos, D., 2002. Mapping excavations and archaeological sites using close range photos, International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, Corfu, September 2-6.

Yeum, C. M., & Dyke, S. J. 2015. Vision-based automated crack detection for bridge inspection. Computer-Aided Civil and Infrastructure Engineering, 30(10), 759-770.