

Marquette University

e-Publications@Marquette

Civil and Environmental Engineering Faculty
Research and Publications

Civil, Construction, and Environmental
Engineering, Department of

11-2020

Determination of Construction Site Elevations Using Drone Technology

Yuhan Jiang
Marquette University

Yong Bai
Marquette University, yong.bai@marquette.edu

Follow this and additional works at: https://epublications.marquette.edu/civengin_fac



Part of the [Civil Engineering Commons](#)

Recommended Citation

Jiang, Yuhan and Bai, Yong, "Determination of Construction Site Elevations Using Drone Technology" (2020). *Civil and Environmental Engineering Faculty Research and Publications*. 273.
https://epublications.marquette.edu/civengin_fac/273

Determination of Construction Site Elevations Using Drone Technology

Yuhan Jiang¹ and Yong Bai²

¹ Department of Civil, Construction and Environmental Engineering, Marquette University, P.O. Box 1881, Milwaukee, WI 53233-1881; PH (414) 378-8632; email: yuhan.jiang@marquette.edu

² Department of Civil, Construction and Environmental Engineering, Marquette University, P.O. Box 1881, Milwaukee, WI 53233-1881; PH (414) 288-6708; email: yong.bai@marquette.edu

ABSTRACT

Image-based 3D-reconstruction techniques, such as drone photogrammetry, have been adopted to various construction operations. The challenge is determining the construction site elevation, which is the vertical distance from the ground to the camera, in real-time. This paper presents the research results of using two frame drone-based ortho-images to estimate the elevation of a construction site. This idea is derived from the stereo vision model for measuring distance of indoor scenes. The spatial resolution of the stereo vision is a positive correlation with its baseline, the distance between two cameras. However, a large baseline stereo cameras system is impossible for a drone to carry. Therefore, a modified stereo vision method is proposed to use a drone's camera to capture low and high ortho-image pairs, which enlarges the baseline, and makes the spatial resolution adjustable for every single construction site. For determining the elevation of each pixel in the image pair, the researchers conducted drone flight planning, image pairs capturing, image transforming, image pixel matching, elevation recovering and modeling. A field experiment was performed for evaluating the accuracy of the proposed method. The success of this research will advance the efficiency of construction operations that heavily depend on elevation information such as the excavation operations and facility layout.

INTRODUCTION

Drone photogrammetry has shown its advantage in various construction related operations, such as building modeling (Chen et al. 2018; Aguilar et al. 2019), pavement modeling and inspection (Inzerillo et al. 2018), construction equipment modeling and tracking (Kim, H. and Kim 2018) and construction site surveying and modeling (Siebert and Teizer 2014; Li and Lu 2018; Moon et al. 2019; Park et al. 2019). These applications rely on a computer vision method, structure from motion (SfM), which uses multiple overlapping images as input, matches scale-invariant feature transform (SIFT)/speeded up robust features (SURF) feature points from the high ratio

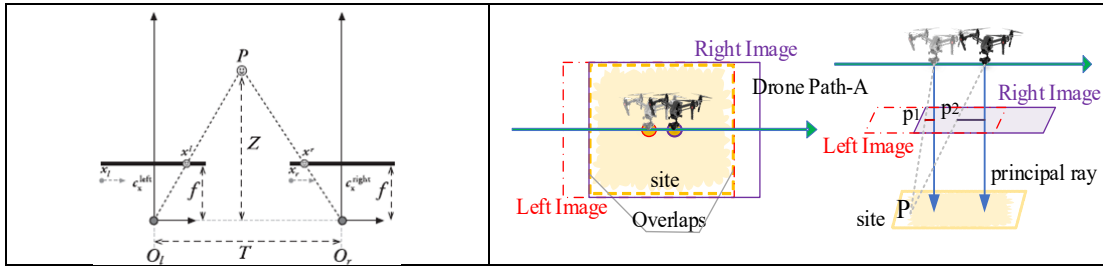
overlapping images and recovers 3D geometrical information for these feature points as the output — 3D point cloud (Lowe 2004; Bay et al. 2008; Nassar et al. 2012; Haur et al. 2018). The advantage of drone-SfM is the commercial softwares, such as *Pix4D*, *Autodesk ReCap Pro*, *Agisoft PhotoScan* and *DroneDeploy*, are available for use. However, this multiple image-based method is a complicated process, which is unable to determine elevations of a construction site in real-time.

Using drone image-based method to determine the elevations of a construction site should be a simpler task than SfM techniques. If the images are captured as top views of the construction site, then this task is equal to determining the vertical distance between the drone (image sensor) and the ground surface. The *stereo camera system* and *RGB-D camera* are two techniques that can be used to measure distance in indoor environments (Sophian et al. 2017; Guo 2018). As the resolution of the *RGB-D camera*'s distance sensor is very small, such as 512×424 pixels for the Kinect V2; and the accuracy of depth measurement gets worse when the distance increases; the *RGB-D camera* is only used for indoor applications, 1.5 m to 3m is an accepted distance range for measurement. Thus, the stereo vision technique is the only potential approach that may be able to extend its application to outdoor environments for determining the distance from the cameras to the ground surface on a construction site. To get the top view of a construction site, the drone's camera is designed to always focus the ground, then an image captured in this condition is named as ortho-image. Therefore, this research problem is defined as determining the vertical distance for each pixel of the overlapping part of two frame ortho-images by triangulation method.

PROPOSED METHOD

Weaknesses of the Standard Stereo Vision

The standard stereo vision method uses the horizontal baseline triangulation to determine distances from cameras to objects (see Figure 1). The captured image pair is two same scale left image and right image, which are in the same plane and perpendicular to the cameras' principal rays. *Eq. 1* indicates objects' distances have a negative relationship with the *Disparity* $= x_l^p - x_r^p$. As the minimum noticeable *Disparity*_{MIN} in an image is 0.5 pixels, so the measurable distance $Z_{MAX} = 2Tf_x$ is dependent on the baseline T (as f_x is fixed).

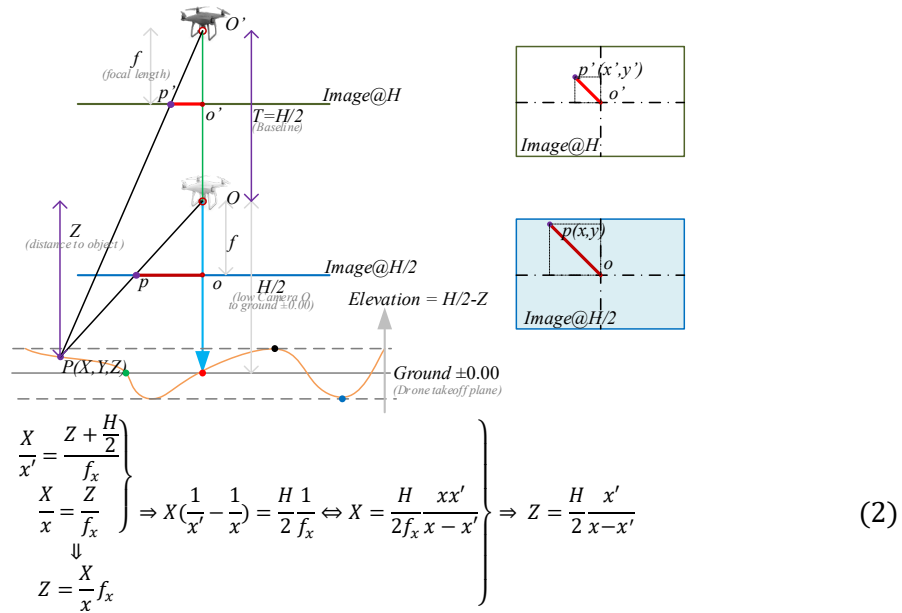


$\left. \begin{aligned} \frac{x_l^p - c_x}{f_x} &= \frac{T_l}{Z} \\ \frac{c_x - x_r^p}{f_x} &= \frac{T_r}{Z} \end{aligned} \right\} \Rightarrow (x_l^p - c_x) + (c_x - x_r^p) = \frac{T_l + T_r}{Z} f_x$ $T_l + T_r = T$ $\Rightarrow \text{Disparity} = x_l^p - x_r^p = \frac{T}{Z} f_x \Leftrightarrow Z = \frac{T}{x_l^p - x_r^p} f_x$ $(x_l^p - x_r^p \neq 0)$	(1)
<p>*P, a target in front of cameras O_l and O_r; $p_l(x_l^p, y_l^p)$ and $p_r(x_r^p, y_r^p)$, image points of P; (c_x, c_y), image point of cameras, ideally is the image center, named as principal point; T, distance between cameras, named as Baseline; f, focal length of cameras sensor, represents in unit: mm; α is the factor to convert sensor size (mm) to image resolution size (pixel); $f_x = \alpha f$, focal length of image, represents in unit: pixel; Z, distance between P and Cameras</p>	
<p>Figure 1. Stereo Camera System, Drone-Based Stereo Vision Model</p>	

However, a drone cannot easily carry a large baseline *stereo camera system*. To simulate the stereo vision method, the baseline T might be extended with drone path-A by capturing the ortho-image pair at two different stations (see Figure 1). That is hard to guarantee these ortho-images are captured in a constant flight height without any off-course movement. Therefore, this drone path-A based stereo vision method is no different to a two-frame SfM method.

Geometry Model of the Modified Stereo Vision

To use a two-frame ortho-image pair to determine the distance from camera to ground surface, the proposed stereo vision method uses the vertical baseline triangulation to determine distances from the low camera to objects (see Figure 2).



$$\text{Similarly, } Z = \frac{Y}{y} f_y \quad \text{and} \quad Y = \frac{H}{2f_y} \frac{yy'}{y-y'} \Rightarrow Z = \frac{H}{2} \frac{y'}{y-y'}$$

$$\text{Thus, } \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{H}{2f_x} \frac{xx'}{x-x'} \\ \frac{H}{2f_y} \frac{yy'}{y-y'} \\ \frac{H}{2} \frac{x'}{x-x'} \text{ or } \frac{H}{2} \frac{y'}{y-y'} \end{bmatrix}, H \in \left[2\alpha f \frac{\max(H_{\text{site}}, W_{\text{site}})}{\max(h_{\text{image}}, w_{\text{image}})}, 2\alpha f \frac{\max(H_{\text{site}}, W_{\text{jobsite}})}{\min(h_{\text{image}}, w_{\text{image}})} \right]$$

* $P(X, Y, Z)$ (in Camera O coordinate) is a target point in the construction site; $p(x, y, \alpha f)$ (in Camera O coordinate) is the image point of P on $\text{Image}@1/2H$; $p'(x', y', \alpha f)$ (in Camera O' coordinate) is the image point of P on $\text{Image}@H$; f_x and f_y are focal length for image in x and y direction, ideally, it has $f_x = f_y = \alpha f$; α is the factor to convert sensor size (mm) to image resolution size (pixel)

Figure 2. Vertical Baseline Stereo Vision Model

In detail, it is proposed to capture the first ortho-image $\text{Image}@H/2$ at the low camera position O with the low flight height $H/2 = \alpha f H_{\text{site}}/h_{\text{image}}$ to cover the entire construction site, and capture the second ortho-image $\text{Image}@H$ at the high camera position O' with the high flight height $H = 2\alpha f H_{\text{site}}/h_{\text{image}}$. Additionally, the vertical baseline $T = H/2$ is the spatial position difference between the low and high camera, and these ortho-images have the same principal point o (o') with a 2:1 scaling relation. Thus, the objects next to the image center have the $x = 1, x' = 0.5$ or $x = -1, x' = -0.5$, and $y = 1, y' = 0.5$ or $y = -1, y' = -0.5$, which are sufficient to be detected with the minimum $\text{Disparity} = x - x'$ or $\text{Disparity} = y - y'$ in 0.5 pixels (Meng et al. 2013). Eq. 2 indicates the objects' distances have a negative relationship with the *Disparity*. If the image point pair $p(x, y, \alpha f)$ on $\text{Image}@H/2$ and $p'(x', y', \alpha f)$ on $\text{Image}@H$ are known, then the target point $P(X, Y, Z)$ can be calculated. Specially, when $x' = x/2, y' = y/2$, it has $Z = H/2$, which means the target point is on the *ground* ± 0.00 . Therefore, target points' elevation (relative to the *ground* ± 0.00) can be calculated by $\text{Elevation} = H/2 - Z$.

Implement Method of Capturing Ortho-image Pair

Drones, such as “DJI Phantom” series, are integrated with the automatic flight control system, GPS, altimeter, gyroscope, inertial measurement unit and other sensors to help them stably hover in midair at the designed position. The 3-axis gimbal (See Figure 3) enhances the camera's stability, when the pitch-axis of the gimbal just arrives at -90° , the camera is facing down to the ground, then the image captured is the ortho-image — the camera's principal ray is perpendicular to the ground plane. In addition, with help from wireless transmission, the relative flight height data is easier to read directly from the drone's remote controller, which set ± 0.00 as the drone takeoff plane. Therefore, the low and high ortho-image pair is easy to capture at the construction site in a very short time without interfering with other construction operations.

The impacts of wind and GPS signal interference cannot be eliminated, which leads to the slightly horizontal displacement or/and rotation happening during the drone moving from the low position to the high position. Therefore, aligning the ortho-image pair to the same center with a slight rotation or/and translation transformation is needed, which makes the principal point o of $\text{Image}@H$ overlap with the principal point o' of $\text{Image}@H/2$.

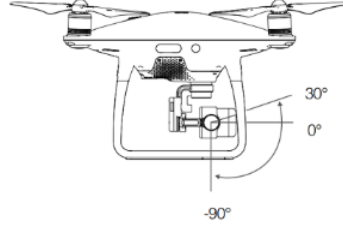


Figure 3. Drone, Camera and Gimbal

Implement Method of Matching Corresponding Points

Each point is the center of a pixel. To find the corresponding point pair in the low and high ortho-image pair is equal to finding the corresponding pixel pair in the ortho-images. Figure 4 explains the procedures to match the pixel pair: 1) converting image pair to grayscale images; 2) building four subsets for $pixel(u, v)$ with 3 of its 8 neighbors, 3) setting the average value of the four subsets as the four features of $pixel(u, v)$; 4) matching the $pixel(u, v)$'s four features with the potential $pixel(u', v')$'s feature by the NCC (normalized cross-correlation) method; 5) determining the subpixel location for $pixel(u', v')$; 6) rotating the image 90, 180 and 270 degrees, and repeat steps 2 to 5; 7) transforming the four pixel-to-subpixel matching results to the same coordinate, and combining the four orientations results.

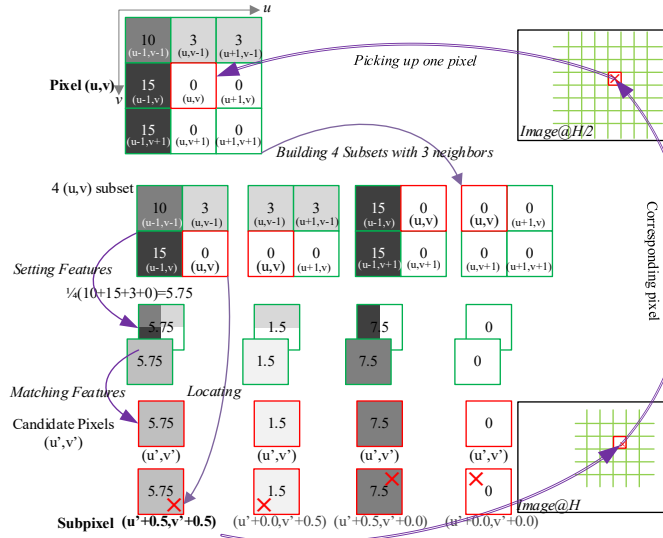


Figure 4. Pixel-to-Subpixel Matching Pipeline

EXPERIMENT

Capturing Ortho-image Pairs

The drone, DJI Phantom Pro 4 V2, took off at Atwater Park in Shorewood, Wisconsin, USA. The drone flew several vertical paths and captured several ortho-images when the controller showed the drone at flight height 10m, 20m, 40m, 60m, 80m and 120m (see Figure 5)

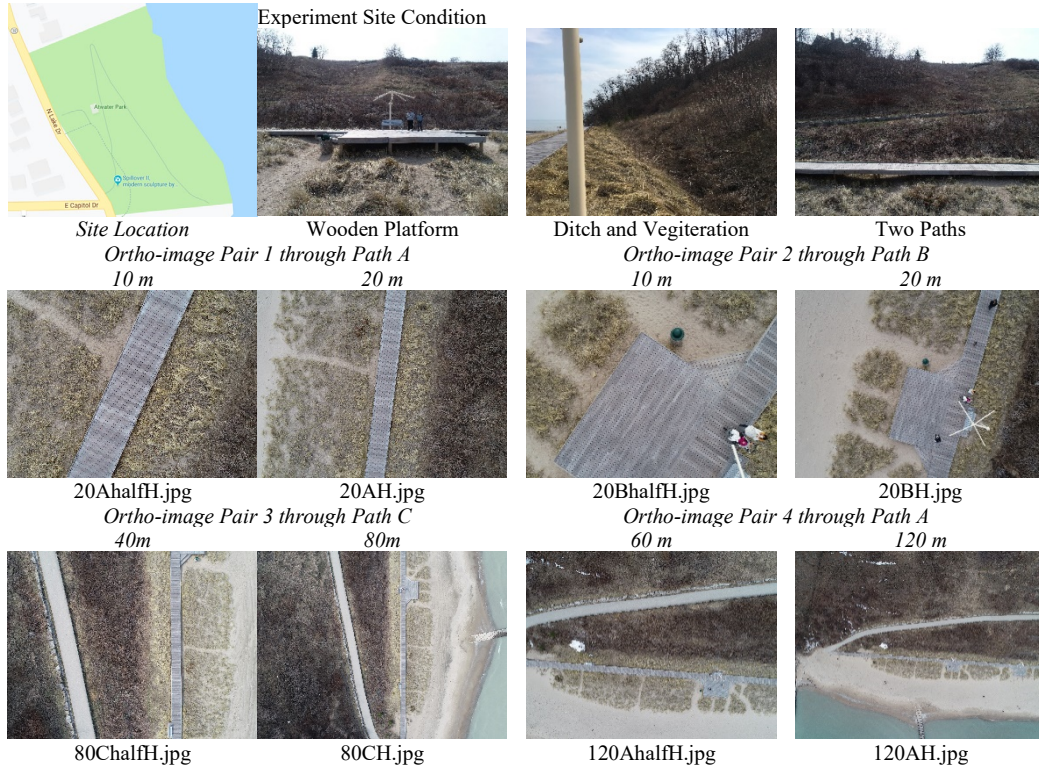


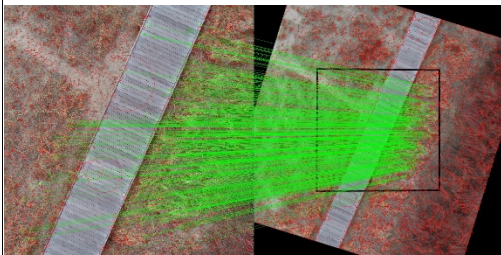
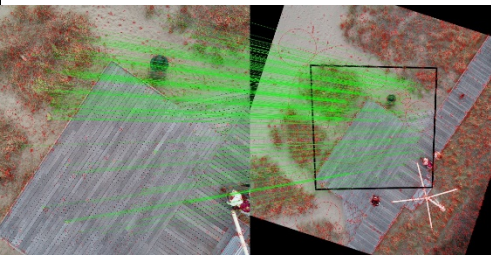
Figure 5. Ortho-image Pair and Site Environment

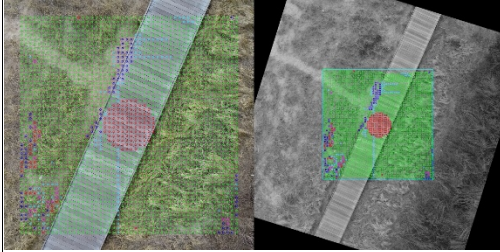
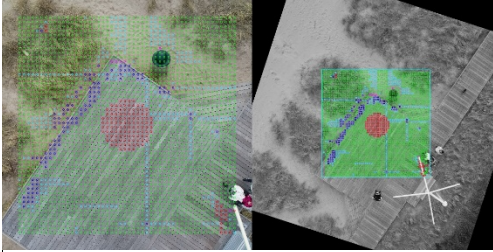
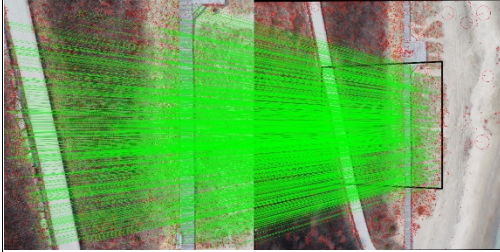
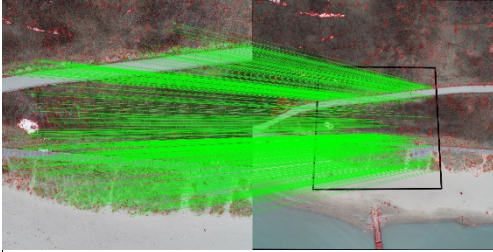
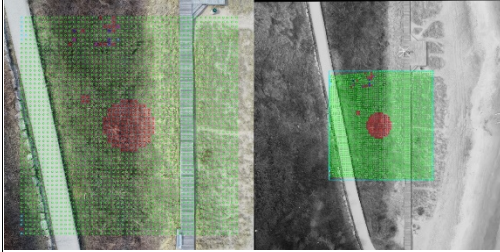
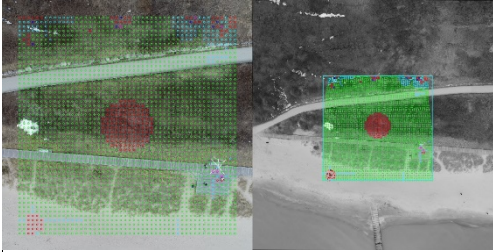
Matching Ortho-image Pairs

The $Image@H$ in the four ortho-image pairs were aligned with $Image@H/2$ using translation and rotation only (without image projection) by image registration (see Table 1). Among these ortho-images, there are visible image rotations in the ortho-image pairs 1 and 2.

The SIFT matching results of these four ortho-image pairs range from 237 to 1512, which is not stable (see Table 1). In addition, the matched keypoints are concentrated in several areas, which is not an even distribution. Furthermore, the edges of the wooden platform are not matched.

Table 1. Ortho-image Pair Matching Results

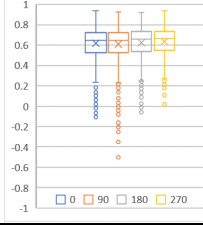
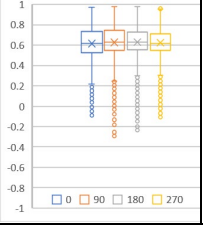
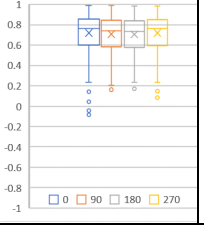
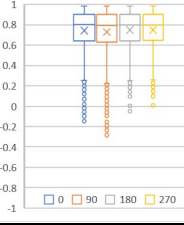
	<i>Ortho-image Pair 1</i>	<i>Ortho-image Pair 2</i>
Translation	(-45.66, 5.42) pixels	(-5.53, -38.20) pixels
Rotation	-17.23 °	-17.99 °
SIFT		
Matched	562 keypoints	237 keypoints

This Paper*		
Matched	2500 Points	2500 Points
	Ortho-image Pair 3	Ortho-image Pair 4
Translation	(1.37, 8.40) pixels	(-2.30, -5.37) pixels
Rotation	0.42 °	-0.20 °
SIFT		
Matched	1512 keypoints	945 keypoints
This Paper*		
Matched	2500 Points	2500 Points

The proposed pixel-to-subpixel matching results use different colors to represent the matching qualities, green > cyan > blue > pink > red (see Table 1). The red center areas were manually changed to bad matching for updating their elevation with adjacent points. Excluding the center area (112 points), the number of bad matching among these four ortho-image pairs ranges from 19 to 56, which accounts for 2.24% of the 2500 selected points at most (see Table 2). Thus, this proposed matching method works better than the SIFT method in this specific task that matches different scales' images. Additionally, the blue stripe area in pairs 1 and 2 are the edge of a wooden path and a platform, where the elevations change sharply, the proposed matching algorithm successfully matched the selected points in these areas. The red area on the left bottom corner of pair 4 is caused by the visitors' movement.

Table 2 also shows the NCC value distribution of four-orientation matching in these four ortho-image pairs (excludes the center area). The lower boundaries for these sixteen boxplots are around 0.2 to 0.3. Thus, using a constant value, such as 0.3, to determine matching quality is reasonable. The median of NCC value is around 0.8 in pairs 3 and 4, which is better than 0.6 in pairs 1 and 2. These statistic results indicate that the pairs 3 and 4 were matched better than pairs 1 and 2 by the proposed matching method. That is the same in SIFT method, and the significant image rotations in pairs 1 and 2 may be the reason.

Table 2. Ortho-image Matching Result Summary

	<i>Pair 1</i>	<i>Pair 2</i>	<i>Pair 3</i>	<i>Pair 4</i>
Total	2500	2500	2500	2500
Good Matching	2364	2369	2371	2332
Green-Four*	2158	1977	2352	2184
Cyan-Three*	100	276	12	135
Blue-Two*	88	110	4	7
Pink-One*	18	6	3	6
Bad Matching	136	131	129	168
Center Area	112	112	112	112
Bad-Center	24	19	17	56
NCC Value Distribution				

* the number of $NCC > 0.3$ among the 0, 90, 180 and 270 orientations

Determining the Elevations

The matched point pairs (x, y) and (x', y') were used to calculate the relative elevation for each pixel. The elevation (range from $-H/4$ to $H/4$ relative to the drone takeoff plane) was represented as an 8-bit grayscale elevation-map (see Figure 6) by $gray_{u,v} = 255 \times (Ele_{u,v} + H/4)/(H/2)$. In pairs 1 and 2, the wooden path and platform are distinguished from the ground. In pairs 3 and 4, the shape of the ground surface is easily noticed by the grayscale value changes in the elevation-map.

Additionally, the selected 2500 *Point* $(X, Y, Ele., R, G, B)$ were imported into the *MeshLab*, and the mesh models were created with the default configuration in MeshLab (see Figure 7).

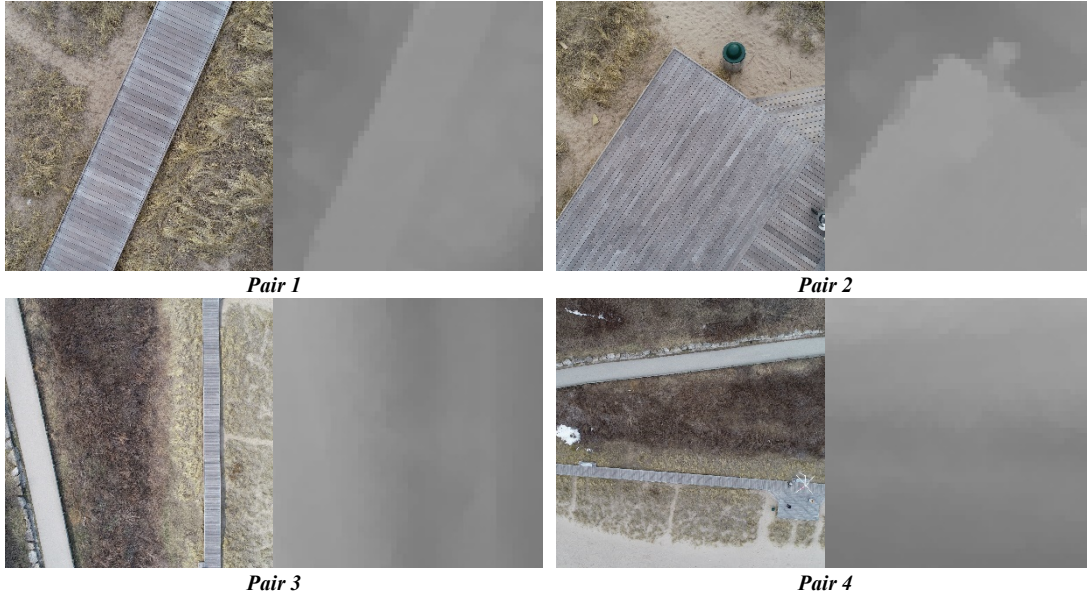


Figure 6. Ortho-image and Elevation-map

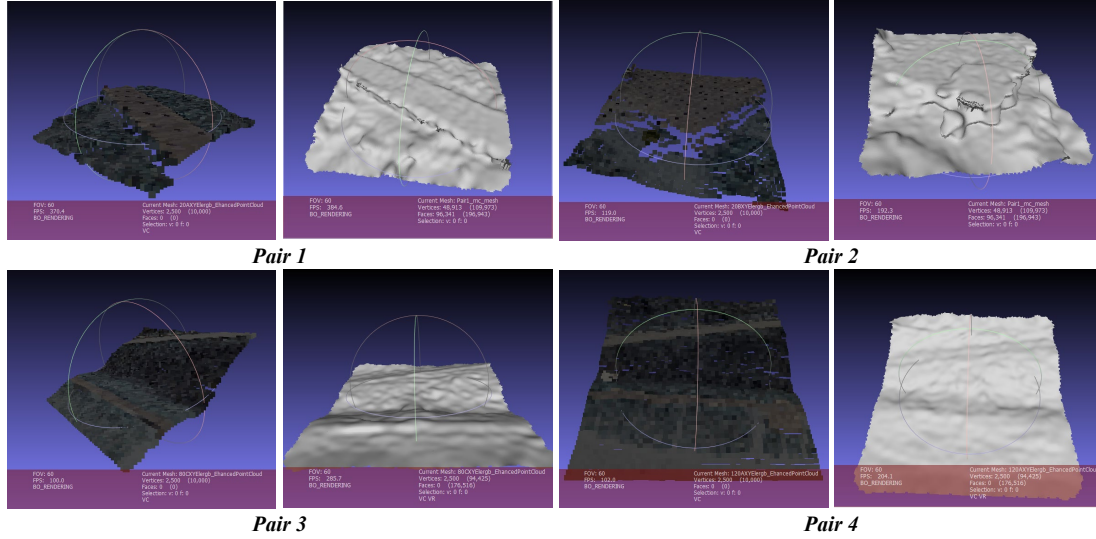


Figure 7. Point Cloud and Mesh Model

Evaluating the Proposed Method

The profiles of the point clouds were used to evaluate the proposed method. With the help of the four orientations matching, the proposed NCC-based pixel matching is sensitive to elevation changes in both x-axis and y-axis direction of ortho-images. In this experiment, the pairs 3 and 4 have an overlap that pair3's x-axis is near to pair 4's y-axis. Thus, pair3's x-axis and pair 4's y-axis should have a similar profile. The experiment result in Figure 8 confirms that the pairs 3 and 4 have the overlap profile from 0 to 17 m, which includes the ditch, wooden path and ground. The shapes of these two profiles from 0 to -17 m are nearly parallel, because this area is a slope.

Pair 1's profile at the y-axis shows the overall wooden path elevation is about 0.85 m, and its profile at the x-axis shows the ground, close to the wooden path, has the elevation around 0.05 m (see Figure 9). Then the wooden path and the ground have about 0.8 m elevation differential. Additionally, five measured vertical distances from the wooden path to the ground is 0.83, 0.82, 0.82, 0.79 and 0.78 m by measuring tape, with the mean 0.808 m. The measure error is - 8 mm (or 0.008 m = 0.80 m - 0.808 m) with the drone flies at 10 meters above than the ground.

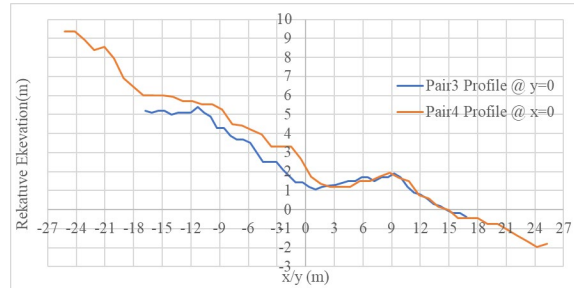


Figure 8. Profile of Pair 3's x-axis and Pair 4's y-axis

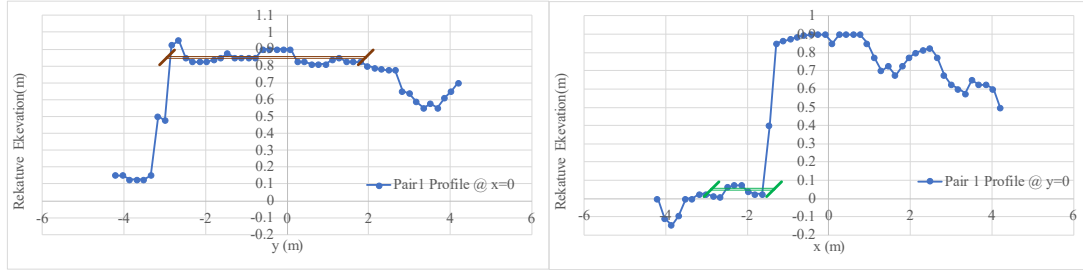


Figure 9. Pair 1 Profiles

CONCLUSION

This paper presents a cheaper, faster and effective 3D reconstruction method, which only needs a drone-based low and high ortho-image pair as the input. The overall procedure includes: 1) using a drone to acquire construction sites' low and high ortho-image pairs, 2) using vertical baseline triangulation model to recover the elevations from the low and high ortho-image pairs, and 3) generating the 3D point cloud and 2D elevation-maps. The proposed method is faster than the standard SfM method. It takes less than 3 minutes to complete the whole procedure for an ortho-image pair (Python 3.7.3, CPU Xeon Gold 5122 @3.6 GHz) with 2500 selected pixels.

In this study, the camera distortion is not considered. This might impact the image feature matching, affecting the horizontal coordinate calculation and horizontal distance measurement (which was not performed in this paper). Further research can undistort the ortho-images with the camera's distortion parameters. Additionally, the experiment only compared the straight-line distance between the model and real-world without aligning the model to the real-world coordinates. Further work can use the ground control points to align the model coordinates to the real-world coordinates. Furthermore, as the experiment only evaluated the 10 m and 20 m other-image pair, further evaluations should be set up to test the accuracy of different drone flight heights, such as 10m and 20m, 60 m and 120 m.

REFERENCES

- Chen, K., Lu, W., Xue, F., Tang, P., and Li, L. H. (2018). "Automatic building information model reconstruction in high-density urban areas: Augmenting multi-source data with architectural knowledge." *Automation in Construction*, 93 22-34.
- Aguilar, R., Noel, M. F., and Ramos, L. F. (2019). "Integration of reverse engineering and non-linear numerical analysis for the seismic assessment of historical adobe buildings." *Automation in Construction*, 98 1-15.
- Inzerillo, L., Di Mino, G., and Roberts, R. (2018). "Image-based 3D reconstruction using traditional and UAV datasets for analysis of road pavement distress." *Automation in Construction*, 96 457-469.
- Kim, H., and Kim, H. (2018). "3D reconstruction of a concrete mixer truck for training object detectors." *Automation in Construction*, 88 23-30.

- Siebert, S., and Teizer, J. (2014). "Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system." *Automation in Construction*, 41 1-14.
- Li, D., and Lu, M. (2018). "Integrating geometric models, site images and GIS based on Google Earth and Keyhole Markup Language." *Automation in Construction*, 89 317-331.
- Moon, D., Chung, S., Kwon, S., Seo, J., and Shin, J. (2019). "Comparison and utilization of point cloud generated from photogrammetry and laser scanning: 3D world model for smart heavy equipment planning." *Automation in Construction*, 98 322-331.
- Park, J., Kim, P., Cho, Y. K., and Kang, J. (2019). "Framework for automated registration of UAV and UGV point clouds using local features in images." *Automation in Construction*, 98 175-182.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.
- Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3), 346-359.
- Nassar, K., and Jung, Y. (2012). "Structure-From-Motion Approach to the Reconstruction of Surfaces for Earthwork Planning." *Journal of Construction Engineering and Project Management*, 2(3), 1-7.
- Sophian, A., Sediono, W., Salahudin, M. R., Shamsuli, M. S. M., and Za'aba, D. Q. A. (2017). Evaluation of 3D-Distance Measurement Accuracy of Stereo-Vision Systems. *International Journal of Applied Engineering Research*, 12(16), 5946-5951.
- Guo, Q., Su, Y., Hu, T., Zhao, X., Wu, F., Li, Y., Liu, J., Chen, L., Xu, G., Lin, G., Zheng, Y., Lin, Y., Mi, X., Fei, L., and Wang, X. (2017). "An integrated UAV-borne lidar system for 3D habitat mapping in three forest ecosystems across China." *International Journal of Remote Sensing*, 38(8-10), 2954-2972.
- Meng, C., Zhou, N., Xue, X., and Jia, Y. (2013). Homography-based depth recovery with descent images. *Machine vision and applications*, 24(5), 1093-1106.
- Haur, C. J., Kuo, L. S., Fu, C. P., Hsu, Y. L., and Heng, C. D. (2018). "Feasibility Study on UAV-assisted Construction Surplus Soil Tracking Control and Management Technique." *IOP Conference Series: Materials Science and Engineering*, 301 12145. <http://doi.org/10.1088/1757-899X/301/1/012145>