



Multi-temporal orthophoto and digital surface model registration produced from UAV imagery over an agricultural field

Taeheon Kim, Jueon Park, Changhui Lee, Yerin Yun, Jinha Jung & Youkyung Han

To cite this article: Taeheon Kim, Jueon Park, Changhui Lee, Yerin Yun, Jinha Jung & Youkyung Han (2022) Multi-temporal orthophoto and digital surface model registration produced from UAV imagery over an agricultural field, Geocarto International, 37:27, 18767-18790, DOI: [10.1080/10106049.2022.2143913](https://doi.org/10.1080/10106049.2022.2143913)

To link to this article: <https://doi.org/10.1080/10106049.2022.2143913>



Published online: 16 Nov 2022.



Submit your article to this journal [↗](#)



Article views: 262



View related articles [↗](#)



View Crossmark data [↗](#)



Multi-temporal orthophoto and digital surface model registration produced from UAV imagery over an agricultural field

Taeheon Kim^a, Jueon Park^a, Changhui Lee^a, Yerin Yun^a, Jinha Jung^b and Youkyung Han^a

^aDepartment of Civil Engineering, Seoul National University of Science and Technology, Seoul, Republic of Korea; ^bLyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA

ABSTRACT

Correcting the three-dimensional geometric error is essential to effectively use the multi-temporal unmanned aerial vehicle (UAV) orthophoto and digital surface model (DSM) acquired from the agricultural field. Although ground control points (GCPs) obtained through field surveys are usually used to calibrate geometrical errors establishing/maintaining GCPs and surveying them in the field are time-consuming and inefficient. Therefore, we propose a simple and efficient methodology to improve the geometric registration of multi-temporal orthophotos and DSMs without GCPs. In the proposed method, coarse to fine image registration is performed first, which corrects severe to slight errors by sequential feature and area-based matching methods. Subsequently, we extract height-invariant regions in multi-temporal DSM pairs, called elevation invariant feature (EIF), using the EIFs to register DSMs by estimating a linear regression model. Various experiments were conducted to analyze the absolute and relative accuracies using ten multi-temporal orthophotos and DSMs, and the robustness of the proposed method was evaluated using data obtained from another site. The experimental results demonstrate that the geometric quality of registered orthophotos and DSMs was significantly improved.

ARTICLE HISTORY

Received 16 March 2022
Accepted 31 October 2022

KEYWORDS

Three-dimensional geometric error; unmanned aerial vehicle (UAV); coarse to fine image registration; elevation invariant feature (EIF)

1. Introduction

For several decades, remote sensing has been a core technology in agricultural applications because it can easily acquire geoinformation and crop data over agricultural land (Bastiaanssen et al. 2000; Mulla 2013). It is crucial to periodically acquire reliable data to monitor crop conditions (Zhang and Kovacs 2012). However, traditional remote sensing using satellite and manned aircraft has limitations in accomplishing the task because of reasons such as clouds, satellite revisit cycles, the high maintenance cost of aircraft, and so on.

Recently, with the expanding accessibility of unmanned aerial vehicle (UAV), acquiring high-quality remote sensing data for agricultural applications is becoming a straightforward task. UAVs can be equipped with very-high-resolution (VHR) sensors to obtain ultra-fine spatial resolution data (Chen et al. 2016; Yun 2017). Moreover, even consumer-grade UAVs are equipped with a global navigation satellite system (GNSS) receiver on the platform to log rough exterior orientation parameters of the acquired images (Nex and Remondino 2014; Uysal et al. 2015). Thanks to the compact size of the platform and minimal operational crew requirements, UAVs can be easily deployed to acquire data on the desired date without a complicated planning process (Honkavaara et al. 2013; Tsouros et al. 2019). For these reasons, UAV is becoming a remote sensing platform for agricultural applications as it can provide research scientists with more reliable and high-quality remote sensing data while overcoming the limitations of the traditional airborne and spaceborne remote sensing platforms.

Structure from motion (SfM) is an image processing technique to generate an orthophoto and a digital surface model (DSM) from raw UAV images (Rokhmana 2015; Gomez and Purdie 2016). One can quickly produce orthophotos and DSMs from UAV images using commercial software such as Pix4d Mapper and Agisoft Metashape. However, even if UAV images are acquired under the same photography conditions for the same area, the geometric properties between multi-temporal UAV orthophoto and DSM differ due to various external factors such as the quality of GNSS sensors, wind conditions, and platform attitude (Xiang and Tian 2011; Zhuo et al. 2017). Accordingly, a three-dimensional (3D) geometric error between the geospatial data products may be inevitable, causing inconsistent results in time-series agricultural applications (Wei et al. 2017; Chebrolu et al. 2018). This study aims to address this issue by developing an algorithm to correct this error without relying on additional configuration.

Traditionally, ground control points (GCP) are often required to correct the 3D geometric error. The 3D geometric error can be minimized by surveying their coordinates and inputting the surveyed coordinates when generating the orthophotos and DSMs. These tasks must be performed whenever UAV imagery is acquired to build accurate time-series UAV orthophotos and DSMs. However, the GCP accuracy is affected by external factors such as the quality of GPS signals, weather, and data acquisition time. In addition, maintaining and surveying GCPs in the field setup is time-consuming and requires significant effort (Han et al. 2019). Furthermore, if the GCP target in the image is not identified, the whole UAV data collection campaign must be repeated (Kim et al. 2019). Accordingly, various attempts have been made to correct the 3D geometric error without GCPs.

Image registration is a method to unify the two-dimensional (2D) coordinates frames between images using tie-points (Choi and Kim 2017; Han 2017; Chang et al. 2018). This methodology is classified into area- and feature-based matching methods according to the procedure of extracting the tie-points, which are core data for performing the registration (Zitová and Flusser 2003; Huo et al. 2012). A UAV orthophoto has a very high spatial resolution such that the features in images are clearly expressed. Thus, Wei et al. (2017) extracted feature points using a Harris detector and performed image registration using the feature points described by the local geometric structure feature and multi-feature descriptors as tie-points. Tsai and Lin (2017) used an accelerated binary robust invariant scalable keypoints (BRISK) algorithm to correct misalignment between UAV orthophotos. Aicardi et al. (2016) performed image registration using tie-points extracted by the block-based scale-invariant feature transform (SIFT) algorithm. Most image registration aims to achieve accurate alignment by improving the performance of the feature-based matching

method. However, crops are dynamic objects that change shape over growing seasons, which makes it challenging to find good tie-points between multi-temporal orthophotos (Raeva et al. 2019). Furthermore, similar spatial patterns are repeated in most agricultural fields, and descriptor vectors of feature points are similar, in turn, increasing the possibility of mismatch (Wang et al. 2014). Accordingly, extracting reliable tie-points between multi-temporal images acquired from the agricultural field is difficult.

Since the image registration focuses on reducing the 2D misalignment between images without GCPs, it is not applicable for reducing the height dissimilarity between DSM. The DSMs, which can be easily produced by images acquired from UAV, is key data for agricultural applications due to its capacity to provide crop heights that can predict the crop growth states or crop types (Kim et al. 2018; Maes and Steppe 2019; Zhang et al. 2021; Li et al. 2022). Accordingly, producing an accurate DSM is important; however, the process of generating the accurate DSM needs GCPs, which requires a very time-consuming task. Gruszczyński et al. (2022) registered the DSM by correcting UAV image-based point clouds using a deep learning network to improve DSM production efficiency. Lee and Oh (2021) modified the DSM by performing matching based on the 3D polynomial model referring to the aerial DSM. Furthermore, Liu et al. (2021) generated an accurate DSM by performing self-calibration based on high-precision synchronization between cameras and the GNSS receiver. These approaches, however, have limitations to apply to the farmland images and DSMs. The deep learning approach, for example, needs numerous training data to achieve reliable results. Furthermore, the deep learning approach has not been extensively explored to correct the height of already generated DSM. Matching based on a 3D polynomial model, assuming that the topographic and object shapes between DSMs are similar, may not work for multitemporal data collected over farmland as the height of crops changes dynamically over the growing season.

Therefore, we propose a simple and efficient approach for improving the geometric correspondence of multi-temporal UAV orthophotos and DSMs acquired over farmland without using GCPs. Coarse to fine image registration that sequentially uses feature- and area-based matching methods is proposed to correct the orthophotos. This registration method aims to correct the orthophoto robustly without GCP independent of time-series changes. Subsequently, the height of the DSMs is corrected using the data considering the measurement of topography within the farmland. The fundamental mechanism of the proposed height registration is to compensate the DSM height to be similar to the ground using the elevation invariant feature (EIF), which is the data of the height-invariant regions between DSMs.

The contributions of this study are: 1) the proposed method effectively registers from severe to slight misalignments of orthophotos without GCPs using the advantages of each feature- and area-based matching method; 2) the proposed method corrects the DSM height without additional data like GCPs using EIFs extracted from the height-invariant regions between DSMs; 3) the proposed method builds accurate time-series orthophotos and DSMs by performing GCP acquisition only once to generate the reference orthophoto and DSM; 4) finally, the performance and accuracy of the proposed method are confirmed through various experiments using the assessment data acquired from field surveying.

2. Methodology

The proposed method aims to minimize 3D geometric error of multi-temporal UAV orthophotos and DSMs without GCPs by referring to the precise 3D coordinates of the reference orthophoto and DSM generated by UAV imagery with GCPs. As shown in

similar performance as the SIFT algorithm, but it can extract tie-points quicker (Zhang et al. 2020; Kim and Han 2021). Furthermore, tie-points are extracted from the SURF algorithm after downsampling the orthophoto to maximize tie-points extraction efficiency.

The SURF algorithm applies the approximated Hessian box filter to the integrated image and extracts a feature point when the corresponding value is the maximum (Bay et al. 2008). The approximated Hessian box filter is generated based on a Gaussian second derivative defined by Equation (1). Then, the orientation of the extracted feature points is assigned, and a descriptor vector is generated to describe them.

$$H(\rho, \sigma) = \begin{vmatrix} L_{xx}(\rho, \sigma) & L_{xy}(\rho, \sigma) \\ L_{xy}(\rho, \sigma) & L_{yy}(\rho, \sigma) \end{vmatrix} \quad (1)$$

where $H(\rho, \sigma)$ is a Hessian matrix with $L_{xx}(\rho, \sigma)$, $L_{xy}(\rho, \sigma)$, and $L_{yy}(\rho, \sigma)$ that are the Gaussian second derivative results associated with the xx , xy , and yy directions, respectively.

The general tie-point extraction method compares the similarity between the described feature point of the reference image and all described feature points of the target image. Then, the most similar feature point pair between reference and target images are extracted as a tie-point (Wang et al. 2014). This method can universally be used when image contains various feature types, such as urban areas. However, the possibility of a mismatch between the feature points increases, when the features in the images are repetitively configured in a particular pattern, such as farmland (Wei et al. 2017). Therefore, we use a searching space to effectively consider spatial characteristics between feature points by limiting the area for comparing similarities. The ground coordinates of the feature points extracted from the reference image are estimated and converted into image coordinates of the target image. Then, the searching space is generated based on locations corresponding to each other. The most similar feature point pair is extracted as a tie-point by evaluating the similarity between the feature points included in the searching space.

To perform coarse image registration, a transformation model should be constructed using the tie-points. In this study, an affine transformation model is used that considers the relationships of scale, nonorthogonality, rotation, and translation between the tie-points (Han et al. 2014), defined as follows:

$$\begin{aligned} X &= a_0 + a_1x + a_2x \\ Y &= b_0 + b_1y + b_2y \end{aligned} \quad (2)$$

where X , Y are the coordinates of reference orthophoto tie-points, x , y are the coordinates of the target orthophoto tie-points, and $a_0, b_0 \dots a_2, b_2$ are the independent affine transformation coefficients.

However, mismatched points that cause distortion when estimating the affine transformation coefficients are included among the extracted tie-points. In this study, we consider the tie-points that are unsuitable for construction in the affine transformation model as mismatched points and eliminate them. The affine transformation model is constructed using the tie-points extracted from each image, and the tie-points of the target image are converted into reference image coordinates. Then, the root-mean-square error (RMSE) determining the distance difference between the tie-points is estimated, and the tie-point with the largest RMSE is considered mismatched and eliminated. These processes are repeated until the RMSE is lower than the threshold value. The tie-points without mismatched points are restored to the coordinates of the raw image with the scale factor used in the previous downscaling. The coarse image registration is performed using the estimated affine transformation model based on the restored tie-points.

2.2. Fine image registration using area-based matching method

The area-based matching method can extract numerous tie-points even when there are no features between images (Bentoutou et al. 2005; Ye et al. 2020). However, it is difficult to extract the tie-points when the misalignment is severe (Paul and Pati 2021). The purpose of fine image registration step is to correct the slight remaining misalignment between images. Hence, the reliability of the extracted tie-points is important. In this study, the fine image registration is performed using the MI algorithm, a representative area-based matching algorithm.

The MI algorithm measures the statistical correlation between the templates of the images. Then, the location with the highest similarity is extracted as a tie-point (Viola and Wells 1997; Gong et al. 2014). The reference and coarsely registered images are segmented by referring to the size of the template. The probability distribution between templates of the reference and the coarsely registered target images is estimated, and the 2D joint histogram is generated by combining the pixel values of the reference and coarsely registered images. Subsequently, the joint probability density function is generated based on the 2D joint histogram. The probability distribution of the reference and coarsely registered target images, the 2D joint histogram, and the joint probability density function are sequentially defined by Equations (3) (4), and (5):

$$p_{Ref}(\alpha) = \sum_{\beta} p_{Ref, Tar}(\alpha, \beta) \quad (3)$$

$$p_{Tar}(\beta) = \sum_{\alpha} p_{Ref, Tar}(\alpha, \beta) \quad (4)$$

$$p_{Ref, Tar}(\alpha, \beta) = \frac{h(\alpha, \beta)}{\sum_{\alpha, \beta} h(\alpha, \beta)} \quad (5)$$

where $p_{Ref}(\alpha)$ and $p_{Tar}(\beta)$ represent the probability distribution of reference and target images, respectively; $h(\alpha, \beta)$ is the combined histogram for reference and target images; α and β represent the axes for reference and target images, respectively; and $p_{Ref, Tar}(\alpha, \beta)$ is the combined probability density function of reference and target images.

The entropy of each image is generated using the joint probability density function and the probability distribution of each image defined by Equations (6) and (7). The combined entropy is calculated using the probability distribution of each image defined by Equation (8). Subsequently, MI is estimated using the entropy and combined entropy for each image defined by Equation (9), and the pixel with the highest MI between images is extracted as a tie-point. The tie-points are extracted from each segmented region by referring to the template size through these mechanisms. The mismatched points removal process is performed in the same way as a step of coarse image registration. Finally, fine image registration is performed by constructing the affine transformation model based on the tie-points without mismatched points.

$$H(Ref) = - \sum_{\alpha} p_{Ref}(\alpha) \log p_{Ref}(\alpha) \quad (6)$$

$$H(Tar) = - \sum_{\beta} p_{Tar}(\beta) \log p_{Tar}(\beta) \quad (7)$$

$$H(Ref, Tar) = - \sum_{\alpha, \beta} p_{Ref, Tar}(\alpha, \beta) \log p_{Ref, Tar}(\alpha, \beta) \quad (8)$$

$$MI(Ref, Tar) = H(Ref) + H(Tar) - H(Ref, Tar) \quad (9)$$

where $H(Ref)$ and $H(Tar)$ are the entropy of reference and target orthophotos, respectively; $H(Ref, Tar)$ denotes the combined entropy of the reference and target orthophotos; and $MI(Ref, Tar)$ represents the MI between reference and target orthophotos.

2.3. DSM registration using EIFs

DSM is 3D data that expresses the height of objects and terrain for use as basic data in various fields. Recently, the DSM can be simply generated by using UAV images, but there is a limit to accurate height estimation without using GCPs. However, acquiring GCPs by performing field surveys and inputting GCPs in generating orthophotos and DSMs are time-consuming tasks. The DSM height without GCPs can be corrected using a polynomial model that minimizes the difference in height values of the same location area between accurate and inaccurate DSMs (Lee and Oh 2021). In this study, we extract the data with a linear relationship between multi-temporal DSMs by considering the characteristics of farmland according to the time-series change to correct the DSM.

The height of the crops in farmland is inconsistent because the shape changes rapidly based on the growth cycle (Gruszczyński et al. 2019). These height values of crops prevent estimating accurate DSM registration model coefficients. It is difficult to use the height value of crops in vegetation areas as data for DSM registration because the height value of crops causes errors. In contrast, ground and artificial structures, which are non-vegetation areas, are regions with invariant height because they do not change shape regardless of seasonal changes. The relative height deviation between DSMs can be identified based on the height value of these non-vegetated areas to estimate the accurate model coefficients for correcting DSMs. Therefore, we defined the height value of the non-vegetation areas between DSMs as EIFs and present in this paper a specific method for extracting EIFs extraction.

The vegetation index (generated by combining some wavelength bands of the image) is effective in extracting and analyzing crops because it expresses the growth status of the crops quantitatively (Na et al. 2018). Therefore, we extract the ground and artificial structure data as EIFs, using the vegetation index based on the optical sensor wavelength. Various vegetation indices can be generated using only red, green, and blue wavelengths. According to studies by Torres-Sánchez et al. (2014) and Yeom et al. (2019), excess green (ExG) developed by Woebbecke et al. (1995) outperforms other RGB-based vegetation indices. Therefore, we use the ExG to extract the ground and artificial structures as EIFs. The ExG is defined as:

$$ExG = 2G_n - R_n - B_n \quad (10)$$

$$R_n = \frac{R}{R + G + B} \quad G_n = \frac{G}{R + G + B} \quad B_n = \frac{B}{R + G + B}$$

where R , G , B are red, green, and blue band reflectance, respectively.

The ExG has high values for vegetation areas and low values for non-vegetation areas. Thus, an additional task must be performed to classify the two areas by estimating the threshold value. In this study, the vegetation and non-vegetation areas are classified based on the threshold value estimated using the Otsu algorithm. However, the vegetation and non-vegetation areas of the binarized ExG image for each image are represented differently because the status and spectral characteristics of the crops differ between the images. Therefore, each binarized ExG is overlapped to extract common non-vegetation areas between the images. Then, the elevation values of the DSMs corresponding to the location

of the common non-vegetation areas between the reference and target images are extracted as EIFs.

The extracted EIFs include outliers because it is impossible to completely distinguish vegetation and non-vegetation areas using the threshold value estimated by the Otsu algorithm. The outliers cause distortion in estimating the DSM registration model coefficients, so they must be eliminated. In this study, the outliers are eliminated using a normalized Z-score estimated based on the difference between the EIFs.

The difference between the reference and target EIFs is estimated as in Equation (11). Subsequently, the mean and standard deviation values are estimated to calculate the normalized Z-score as defined in Equation (12). The height invariant is robust as the normalized Z-score value is closer to 0. Therefore, EIFs with a normalized Z-score value corresponding to the lower 30% were only used for the DSM registration by eliminating others.

$$D_{EIFs} = |EIFs_{Ref} - EIFs_{Tar}| \quad (11)$$

$$N_{EIFs} = \sqrt{\left(\frac{D_{EIFs} - \mu_{D_{EIFs}}}{\sigma_{D_{EIFs}}}\right)^2} \quad (12)$$

where D_{EIFs} is the EIFs differences; $EIFs_{Ref}$ and $EIFs_{Tar}$ are the EIFs of the reference and target DSMs, respectively; N_{EIFs} represents the normalized Z-score; $\mu_{D_{EIFs}}$ is the mean of the EIFs difference; and $\sigma_{D_{EIFs}}$ represents the standard deviation of the EIF difference.

The EIFs were used to register the DSM using a linear regression model. The offset and gain of the model are defined by Equations (13) and (14), respectively:

$$Y' = aX + b \quad (13)$$

$$a = \frac{\sigma_{Ref}^{EIFs}}{\sigma_{Tar}^{EIFs}}, \quad b = \mu_{Ref}^{EIFs} - a\mu_{Tar}^{EIFs} \quad (14)$$

where Y' represents the corrected target DSM; X is the target DSM; a and b are registration coefficients of the gain and offset, respectively; σ_{Ref}^{EIFs} and σ_{Tar}^{EIFs} are the EIF standard deviations of reference and target DSMs; and μ_{Ref}^{EIFs} and μ_{Tar}^{EIFs} represents the EIF means of the reference and target DSMs.

3. Experimental data set description

The farmland located in Sangju City, South Korea, was selected as the experimental site to evaluate the performance and accuracy of the proposed method (Figure 2). The site comprises field crops such as peppers, soybeans, sesame, perillas, and rice paddies.

Multi-temporal UAV images used in the experiment were acquired by mounting the optical sensor Zenmuse X4S on the Inspire-2, a rotorcraft UAV. UAV images were acquired through 10 flights over approximately four months, from June 21, 2019, to October 4, 2019. The image acquisition conditions were similar, at an acquired time around 10–11 am, an altitude of 30 m, a vertical and horizontal overlap of 80%, and a flight speed of 5–7 m/s. 10 orthophotos and DSMs were produced using Metashape software based on the acquired multi-temporal UAV images. Moreover, some orthophotos and DSMs were generated by inputting GCPs acquired by a field survey. The specifications of the UAV, optical sensor, and GNSS receiver used in this study are presented in Table 1. The specifications of the multi-temporal orthophotos and DSMs are shown in Table 2.

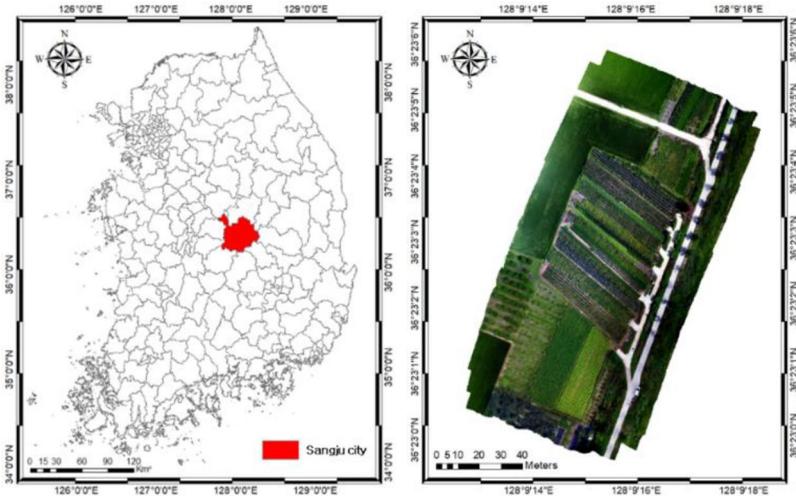


Figure 2. Experimental site: Sangju City, Gyeongsangbuk-do. South Korea.

Table 1. Specifications of equipment to acquire experimental data.

Inspire-2		Zenmuse X4S		Trimble R8s	
					
Weight	3,440 g	Sensor	Optical	Channels	440
Flight altitude	≤ 2500 m	Resolution	5472 × 3648	Satellite signals	GPS: L1C/A, L1C, L2C, L2E, L5
Flight time	27 min	Focal Length	8.8 mm	VRS precision	Horizontal: 8 mm + 0.5 ppm RMS Vertical: 15 mm + 0.5 ppm RMS
Speed	≤ 94 km/h	Weight	253 g	Static Precision	Horizontal: 3 mm + 0.1 ppm RMS Vertical: 3.5 mm + 0.4 ppm RMS

Table 2. Specification of time-series orthophotos and DSMs.

Acquisition date	Scene size (unit: pixel)	Spatial resolution
2019-06-21	9142 × 12037	Approximately 1 cm
2019-07-09 (with GCPs)	9427 × 12606	
2019-07-22	9409 × 12177	
2019-08-01 (with GCPs)	9721 × 12426	
2019-08-14 (with GCPs)	9355 × 12345	
2019-08-20 (with GCPs)	9500 × 12612	
2019-09-06	9229 × 12485	
2019-09-10	10262 × 13595	
2019-09-23	9968 × 13469	
2019-10-04	10301 × 13697	

An accuracy evaluation was performed on datasets with GCPs survey results. Among the surveyed GCPs, a half was used for modelling the orthophoto and DSM, and the remaining GCPs were used as checkpoints to evaluate the performance of the proposed algorithm (Figure 3). The accuracy of each orthophoto and DSM is presented in Table 3, derived by estimating the RMSE in X, Y, and Z directions. Consequently, the UAV

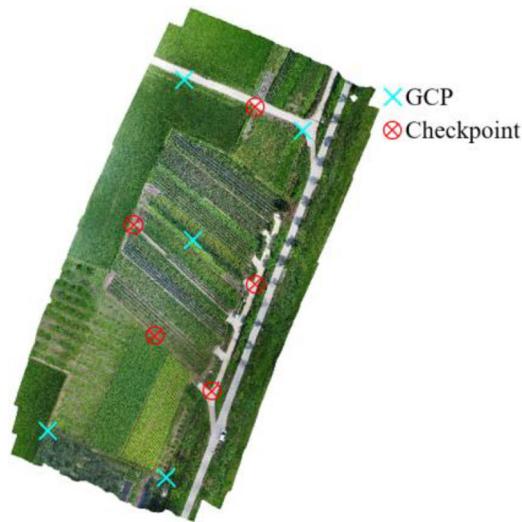


Figure 3. Distribution of GCPs and checkpoints used in experiments.

Table 3. Accuracy of multi-temporal orthophotos and DSMs with GCPs (unit: cm).

Acquisition date	X error	Y error	Z error	Total error
2019-07-09	1.308	1.364	6.149	6.433
2019-08-01	1.081	0.934	5.392	5.515
2019-08-14	0.691	1.140	11.180	11.259
2019-08-20	1.154	0.540	10.781	10.856

orthophoto and DSM acquired on 2019-08-01 in which the total error was the lowest were selected as reference data.

Radiometric calibration for the optical sensor is performed using reflectance panels with intrinsic reflectance. The reflectance provided in the reflectance panels is linearly related to the brightness value of the image. Thus, the radiometric calibration can be performed by constructing a linear regression model (Wang and Myint 2015). In this study, four reflectance panels with surface reflectance obtained using a Perkin-Elmer Lambda 1050 spectroradiometer are used for radiometric calibration.

4. Experimental results and analysis

To confirm the performance of the proposed method, we performed an analysis of the results with one representative dataset, and the related description is given in Section 4.1. In Section 4.2, an absolute accuracy assessment was carried out based on UAV datasets collected together with GCP surveying data. Finally, relative accuracy was assessed for all the multi-temporal orthophotos and DSMs in Section 4.3.

4.1. Results of orthophoto and DSM registration

In this section, we visually accessed the results of registered orthophoto and DSM produced by applying the proposed method. Specifically, we analyzed the distribution of tie-points and identified the registration results by generating mosaic images before and after registration. Moreover, the extraction results of the EIFs were analyzed, and the registered DSM was expressed as point clouds for visual analysis.

To perform coarse image registration, tie-points were extracted using the SURF algorithm and searching space, and mismatched points were eliminated. The size of the searching space was set to 400×400 pixels, determined by empirical analysis on the experimental sites. Most of the tie-points were extracted from clearly identified areas such as the road, farmland division line, and boundary (Figure 4a). Subsequently, we extracted tie-points for fine registration using the MI algorithm while moving the 400×400 size template at 200-pixel intervals. Mismatched points were removed using the same method as the coarse registration step. As a result, the tie-points were distributed on the road and farmland between images (Figure 4b). The tie-points for fine registration tend to be located in areas with similar spectral characteristics between images. These results indicate the characteristics of each matching method.

Mosaic images before and after performing image registration were generated for visual analysis. Furthermore, a comparative analysis was performed to confirm the differentiation of the proposed method compared with other methods. Accordingly, we generated mosaic images of the registration results produced independently with the SURF and MI algorithms (Figure 5). The reference orthophoto is expressed as Red, Green, and Blue (RGB), and the target orthophoto (2019-06-21) is expressed as Blue, Green, and Red (BGR), in the mosaic images. In the mosaic image before registration, the roads and objects were misaligned between the images (Figure 5a). The SURF algorithm and the proposed method successfully aligned the discrepancy between images (Figure 6b and d). However, the misalignment was significantly more severe than the raw image when using the MI algorithm (Figure 5c). Additionally, we evaluated the registration results between the proposed method and the SURF method by visually checking parts of the mosaic images as shown in Figure 6. Although the overall misalignment was minimized when using the SURF algorithm, the slight misalignment remained (Figure 6a). Conversely, the proposed method reduced the remaining slight misalignment (Figure 6b). Therefore, the proposed method corrected the misalignment sequentially using the feature-and area-based matching method.

After carrying out registration, ExG image was generated by combining the optical bands of each image to correct the DSM height without GCP. Each ExG image was binarized into vegetation and non-vegetation areas based on the threshold value estimated using the Otsu algorithm. Then, each binarized ExG image was overlapped, and the height values corresponding to the common non-vegetation areas between images were

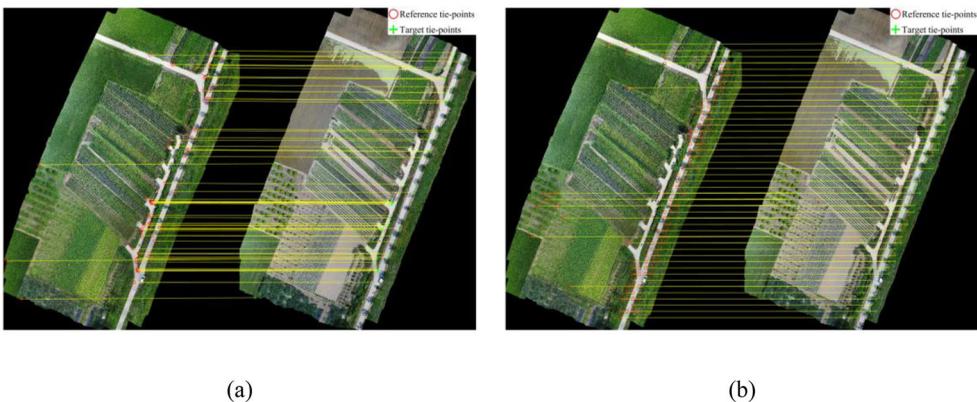


Figure 4. Results of extraction of tie-points using proposed method: (a) Coarse image registration and (b) Fine image registration.

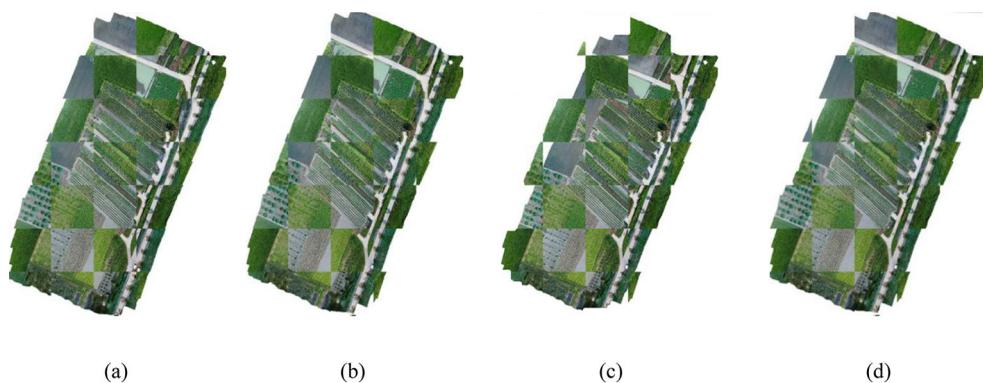


Figure 5. Mosaic images of image registration results using SURF, MI, and proposed methods: (a) Raw image, (b) SURF algorithm, (c) MI algorithm and (d) Proposed method.

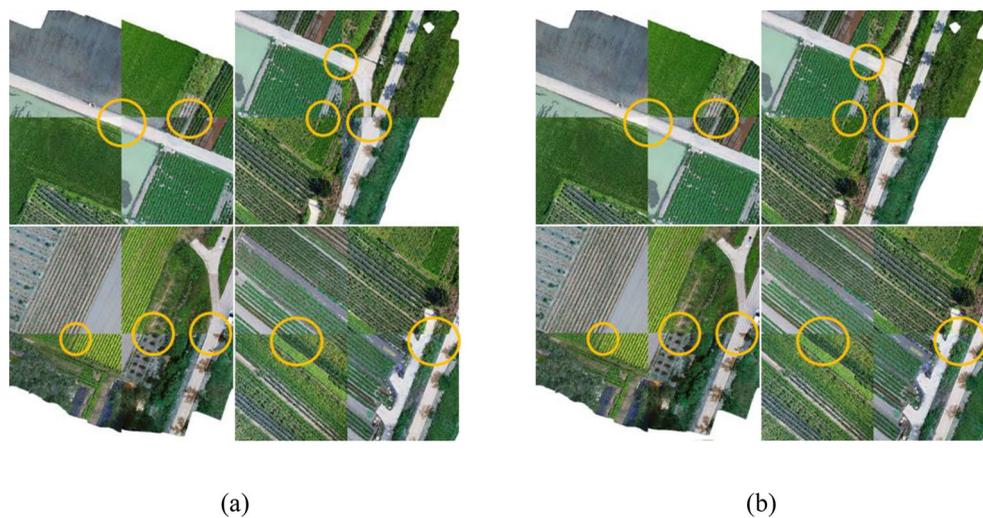


Figure 6. Magnified examples of mosaic images: (a) SURF algorithm and (b) Proposed method.

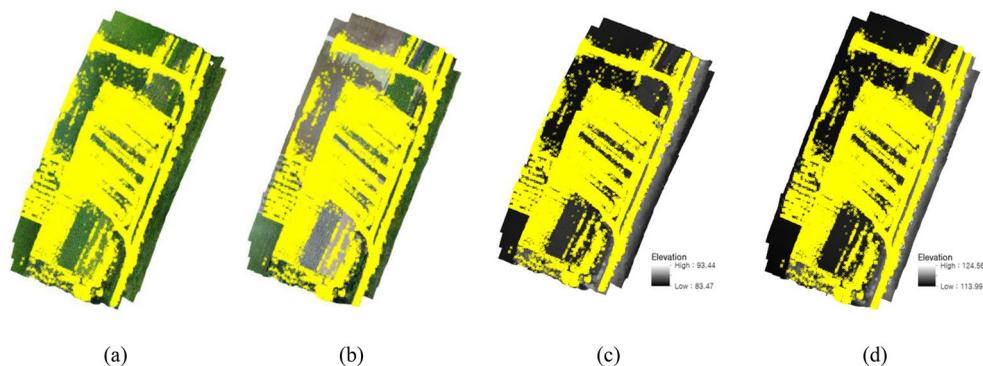


Figure 7. Results of extracting EIFs: (a) Reference image, (b) Target image, (c) Reference DSM, and (d) Target DSM.

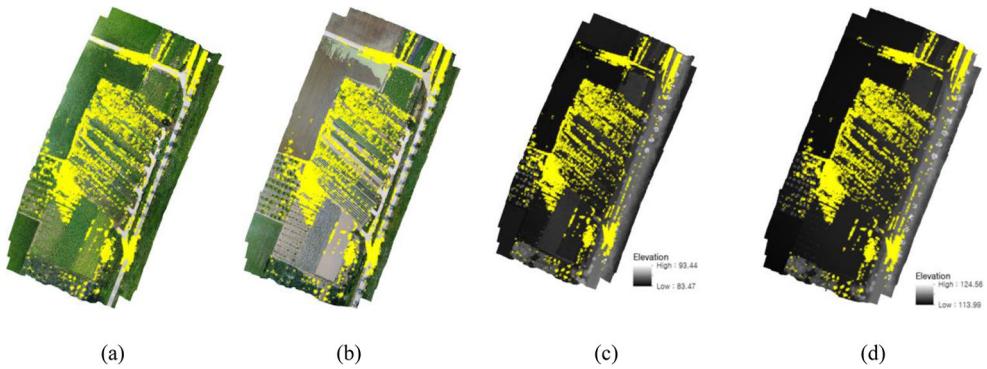


Figure 8. Results of extracting EIFs after eliminating outliers: (a) Reference image, (b) Target image, (c) Reference DSM, and (d) Target DSM.

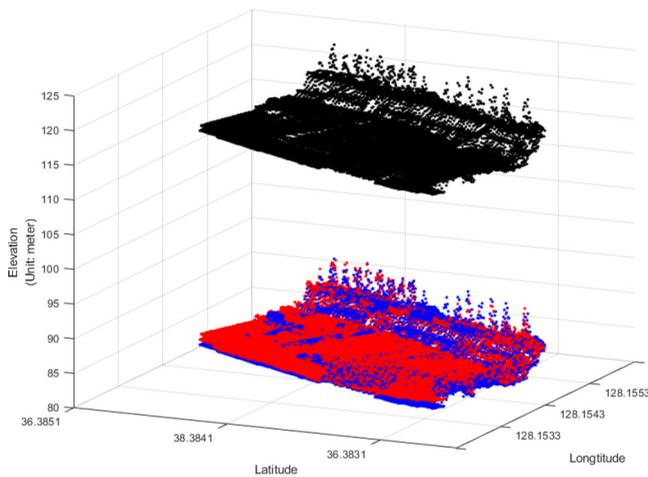


Figure 9. Point cloud of each DSM in 3D space. Point clouds of reference DSM are red, and point clouds of raw and corrected DSMs are black and blue, respectively.

extracted as EIFs. As a result, most EIFs were extracted from non-vegetation areas as shown in Figure 7. However, some EIFs extracted from the vegetation area were still included. These outliers were removed by applying normalized Z-score values. As presented in Figure 8, EIFs after removing outliers were mainly remained on road and ground areas.

Subsequently, target DSMs were registered by calculating a linear regression model based on the EIFs. To visually assess the DSM registration results, each DSM was converted into point clouds and expressed in 3D space as shown in Figure 9. The point clouds of reference DSM are expressed as red, and the point clouds of the raw and corrected DSMs are black and blue, respectively. The point clouds of the raw DSM were generated at approximately 115–125 m in height. In contrast, the point clouds of the corrected DSM were generated at approximately 83–95 m, similar to the height of the reference DSM. The proposed method corrected the terrain height and the height of objects such as crops, and trees, similar to the reference DSM. Therefore, the EIFs extracted from non-vegetation areas reflected the relative elevation deviation between the reference and target DSMs.

4.2. Accuracy assessment of corrected orthophotos and DSMs

In this section, the absolute accuracy of the registered orthophotos and DSMs using the proposed method are evaluated. To this end, datasets with available checkpoints acquired using a GNSS receiver were used. The checkpoints exist only in four datasets among the experimental data in Table 2. Therefore, the absolute accuracy assessment was performed on three datasets except for the reference data by estimating the RMSE based on the checkpoints.

Each experimental orthophoto was reconstructed for five cases to analyze the accuracy of the corrected orthophoto: 1) orthophoto generated with GCPs (baseline); 2) orthophoto generated without GCPs; 3) registered orthophoto using the SURF algorithm; 4) registered orthophoto using the MI algorithm; 5) registered orthophoto using the proposed method. The absolute accuracy assessment was focused on a comparative analysis based on the reconstructed orthophotos in these five cases.

The estimated RMSEs based on the checkpoints of each orthophoto are given in Table 4. The average RMSEs of the orthophoto generated with GCPs were estimated as 0.011 and 0.010 m in the X and Y directions, respectively. Because the orthophoto without GCPs was generated using the coordinate of the UAV GNSS receiver, the average RMSEs were estimated as 0.420 and 1.181 m in the X and Y directions, respectively. In the registered orthophoto using the SURF algorithm, the RMSEs in both directions were estimated as lower than the orthophoto generated without GCPs. The average RMSEs were estimated as 0.075 and 0.055 m, which decreased by approximately 5.6 and 21 times in the X and Y directions compared with the orthophoto without GCPs, respectively. Conversely, the RMSEs in both directions of the registered orthophoto using the MI algorithm was estimated as significantly larger than the orthophoto without GCPs. In the corrected orthophoto using the proposed method, the positional accuracy was improved compared with the case of using the SURF algorithm; the average RMSE was estimated to be 0.024 and 0.037 m in the X and Y directions. These results demonstrate that the proposed method can correct an orthophoto generated without GCPs with a positional accuracy similar to one generated with GCPs. In addition, the superiority of the proposed technique can be visually confirmed as shown in Figure 10.

Each experimental DSM was reconstructed for three cases to evaluate the accuracy of the corrected DSM: 1) DSM generated with GCPs (baseline); 2) DSM generated without GCPs; 3) registered DSM using the proposed method. The accuracy assessment for DSMs focused on whether the proposed method can improve the elevation of the DSM similar to the DSM generated with GCPs.

The RMSEs based on the checkpoint elevations of each formation of DSM are summarized in Table 5. The RMSEs of the DSMs with GCPs for each acquisition date were estimated between 0.061 and 0.112 m. The DSMs without GCPs of the experimental data were generated based on the relative elevation acquired from the UAV GNSS receiver. Thus, the average RMSE was estimated as 30.433 m. The elevation RMSE for each acquisition date was estimated to be larger than the RMSEs in the X and Y directions. In the

Table 4. Results of estimating the RMSE of each orthophoto using the checkpoints (Unit: m).

Acquisition date	Orthophoto with GCPs		Orthophoto without GCPs		SURF algorithm		MI algorithm		Proposed method	
	X	Y	X	Y	X	Y	X	Y	X	Y
2019-07-09	0.013	0.014	0.415	1.186	0.036	0.076	1.588	6.550	0.036	0.050
2019-08-14	0.007	0.011	0.444	1.372	0.051	0.034	2.185	9.419	0.014	0.025
2019-08-20	0.012	0.005	0.402	0.986	0.139	0.055	1.384	5.478	0.021	0.035
Average	0.011	0.010	0.420	1.181	0.075	0.055	1.719	7.149	0.024	0.037

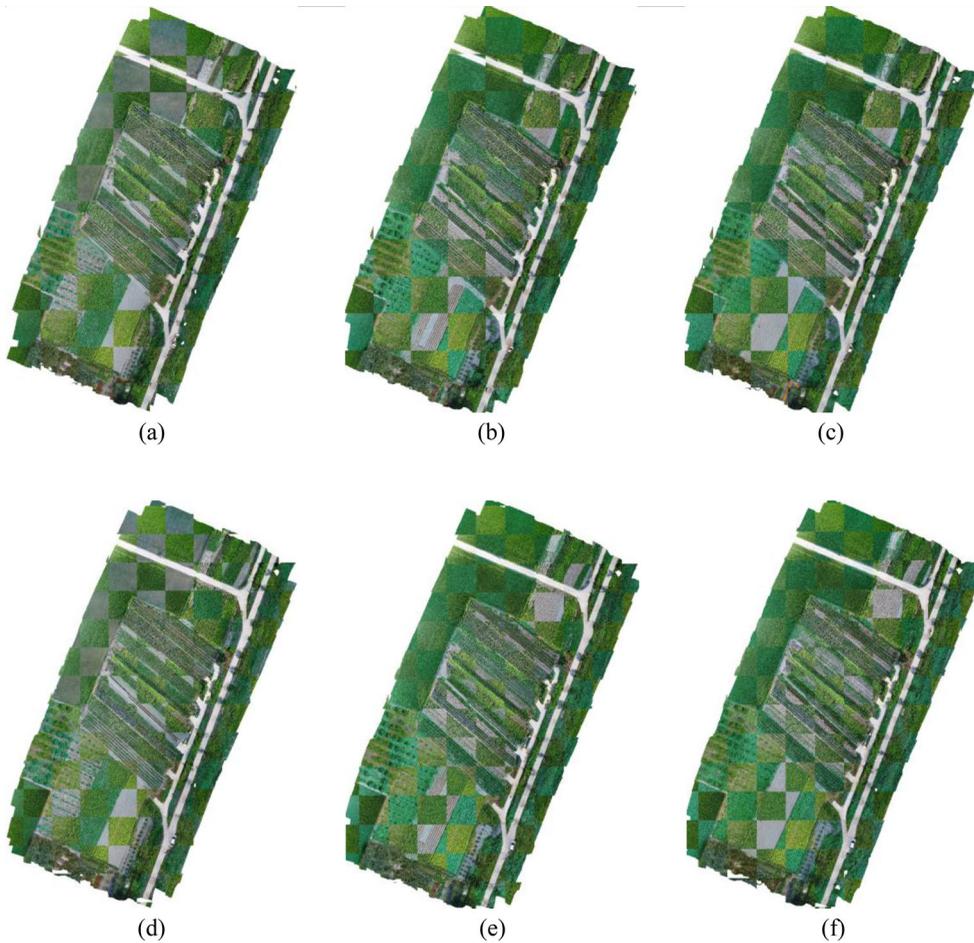


Figure 10. Mosaic images of orthophotos registered by proposed method: (a) Before registration of 2019-07-09 orthophoto, (b) Before registration of 2019-08-14 orthophoto, (c) Before registration of 2019-08-20 orthophoto, (d) After registration of 2019-07-09 orthophoto, (e) After registration of 2019-08-14 orthophoto, and (f) After registration of 2019-08-20 orthophoto.

Table 5. Estimated RMSEs of each formation DSM using checkpoints (Unit: m).

Acquisition date	DSM with GCPs	DSM without GCPs	Corrected DSM
2019-07-09	0.061	8.510	0.214
2019-08-14	0.112	55.852	0.156
2019-08-20	0.107	26.937	0.082
Average	0.093	30.433	0.151

corrected DSM, the RMSE of each acquisition date significantly decreased compared with the DSM without GCPs. Consequently, the average RMSE was estimated as 0.151 m, which is an improvement of more than 30 m compared with the average RMSE of the DSMs without GCPs. Comprehensively, these results demonstrate that the elevation of the DSM without GCP can be corrected similarly to a DSM with GCPs using the proposed method. Furthermore, the elevation deviation between the multi-temporal DSMs was minimized as shown in [Figure 11](#), because the elevation of the DSM without GCPs was corrected similarly to the absolute ground elevation.

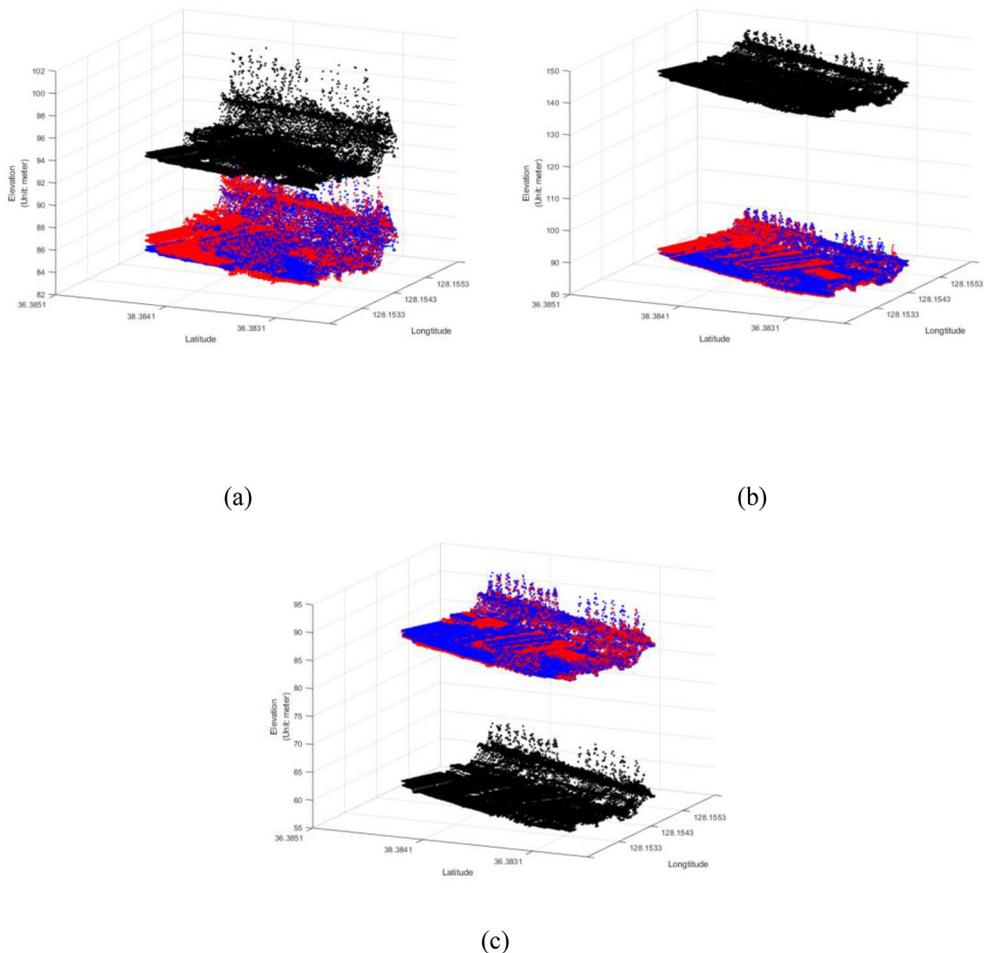


Figure 11. Point clouds of multi-temporal DSMs: (a) 2019-07-09 DSM, (b) 2019-08-14 DSM, and (c) 2019-08-20 DSM. Point clouds of reference DSM are red, and point clouds of raw and corrected DSMs are expressed as black and blue, respectively.

4.3. Results of time-series orthophotos and DSMs

Accurate time-series orthophotos and DSMs are required to acquire data based on the growing season of crops for agricultural applications. Based on the previous assessment results, the geometric accuracy of orthophotos and DSMs generated by applying the proposed method are similar to those generated with GCPs. Therefore, we registered to generate the time-series orthophotos and DSMs listed in Table 2 using the proposed method.

Unfortunately, not all datasets have GCPs with GNSS surveying. Additional reference data are needed to evaluate the accuracy of registered multi-temporal orthophotos and DSMs. Therefore, we selected seven relative checkpoints (RCP) for accuracy assessment by visually identifying the invariant regions between the reference and the multi-temporal data as shown in Figure 12. All RCPs were extracted from the ground or road where the location and height are the same regardless of the acquisition conditions.

The corrected multi-temporal orthophotos using the proposed method were analyzed by comparing the results of the SURF and MI algorithms. Besides, RMSEs before and

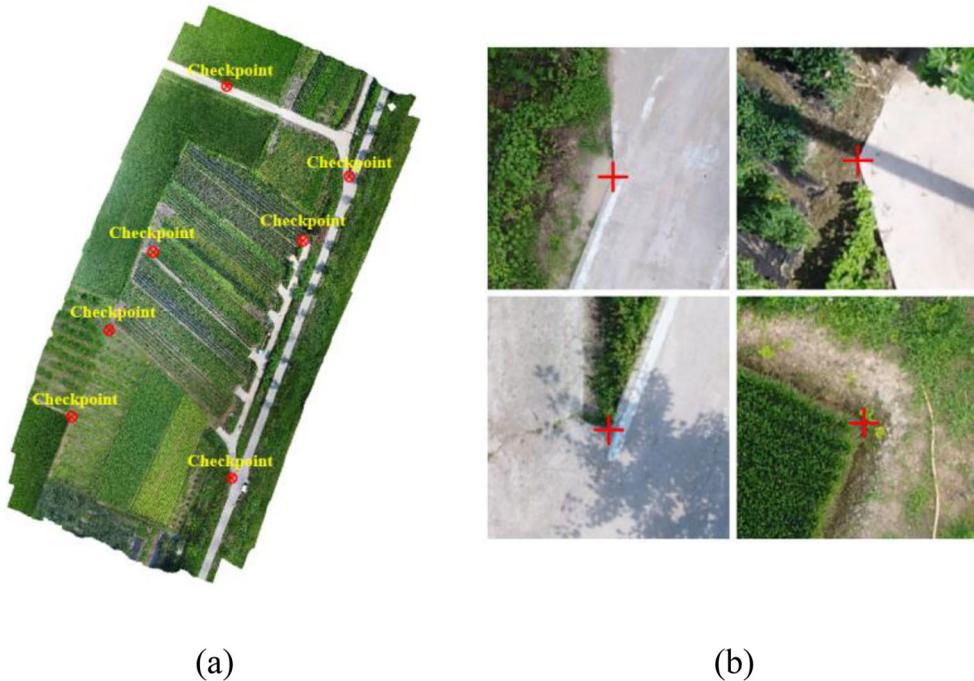


Figure 12. Distribution of selected RCP: (a) Location and (b) Examples.

Table 6. RMSEs of image registration results (Unit: m).

Acquisition date	without GCPs (no registration)	SURF algorithm	MI algorithm	Proposed method
2019-06-21	0.822	0.541	3.276	0.039
2019-07-22	0.921	0.241	2.438	0.019
2019-09-06	1.132	0.314	3.670	0.055
2019-09-10	0.961	0.516	3.356	0.027
2019-09-23	0.918	1.906	2.094	0.025
2019-10-04	0.744	1.647	2.503	0.044
Average	0.916	0.861	2.889	0.034

after registration were estimated by considering both X and Y directions based on RCPs to confirm the accuracy of large amounts of data.

Table 6 presents the RMSEs for the registration results of each method. The average RMSE of the orthophoto without GCPs was estimated as 0.916m. With the SURF algorithm, the misalignment between reference and most orthophotos was minimized, such that the average RMSE was reduced to 0.861 m. However, the misalignment between reference and 2019-09-23 and 2019-10-04 orthophotos became severe, increasing the RMSE. When using the MI algorithm, the misalignment of the corrected orthophoto became more severe than the orthophoto without GCPs, and the average RMSE was estimated as 2.889 m. The RMSE of the proposed method was estimated to be the lowest value among the results. Hence, the average RMSE was estimated as 0.034 m. The proposed method corrected the misalignment in the orthophotos (2019-09-23 and 2019-10-14) for which the SURF algorithm failed. Comprehensively, these results demonstrate that an accurate time-series orthophoto for agricultural applications was successfully generated using the proposed method.

The accuracy of corrected multi-temporal DSM was evaluated as follows: 1) box plot of each DSM and 2) RMSE based on RCP. A box plot is generated by estimating the

minimum, maximum, median, first quartile, and third quartile values of the data and expresses the statistical characteristics of the various data in one space. Because a box plot is generated using the height of the DSM before and after DSM registration, the height change of the DSM can be analyzed intuitively.

The box plots generated based on the height value of the DSM before and after DSM registration are depicted in Figure 13. The box plot (indicated by the yellow box) illustrates the reference DSM, and the red dotted line represents the box plot maximum-minimum value range of the reference DSM. Before DSM registration, the height ranges of the DSM box plots of 2019-06-21 and 2019-07-22 are higher than the reference DSM box plot, whereas those of the other DSM box plots are lower than that of the reference DSM (Figure 13a). After conducting DSM registration, the height range of corrected DSM box plots was expressed similarly to the reference DSM box plot (Figure 13b).

The proposed method corrected the height of the target DSM even when the relative height deviation was significant. To quantitatively evaluate these results, the RMSEs before and after DSM registration were estimated using RCPs, summarized in Table 7. The RMSEs for the DSM without GCPs were estimated to be a minimum of 16.322 m and a maximum of 46.924 m, and the average RMSE was estimated as 34.866 m. The height error of DSM without GCPs was more sensitive to the initial value than orthophoto without GCPs. Hence, the RMSEs were estimated to be much larger than the 2D error in the X and Y directions. The RMSEs of the corrected DSM using the proposed method were estimated to be a minimum of 0.151 m and a maximum of 0.484 m, and the average RMSE was estimated to be 0.251 m. Based on these results, the accuracy of the corrected DSM improved compared with the DSM without registration.

5. Discussion

In this study, we proposed a methodology to build consistent time-series orthophotos and DSMs produced from UAV imagery over farmland. The geometric and height dissimilarity between reference and multi-temporal orthophotos and DSMs were minimized by the proposed method without using GCP. In this section, a detailed analysis with respect to the proposed method focusing both on the image registration and DSM registration was performed. Moreover, experiments were conducted on an additional study site to verify the robustness of the method.

In the registration results described in Sections 4.1–4.3, the misalignment became more severe than the raw image while using only the MI algorithm due to the mechanism of the area-based matching method. The template should be overlapped at similar locations

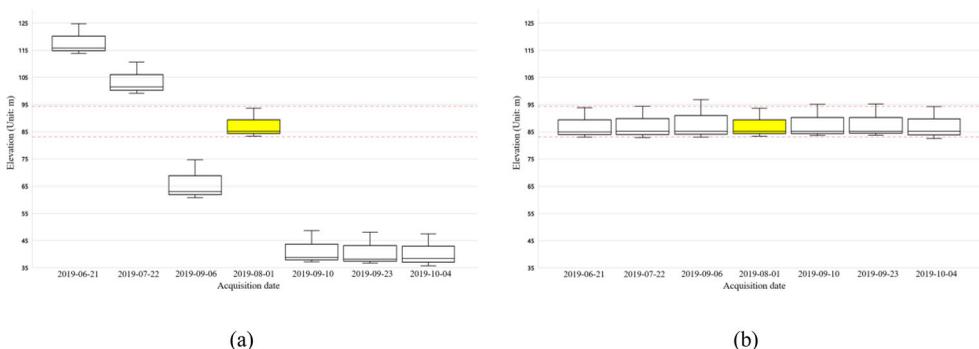


Figure 13. Box plots based on height of each DSM: (a) Before DSM registration and (b) After DSM registration.

Table 7. RMSEs of DSM registration results (Unit: m).

Acquisition date	DSM without GCPs	Registered DSM
2019-06-21	30.905	0.151
2019-07-22	16.322	0.282
2019-09-06	22.078	0.484
2019-09-10	46.321	0.171
2019-09-23	46.924	0.241
2019-10-04	46.645	0.180
Average	34.866	0.251

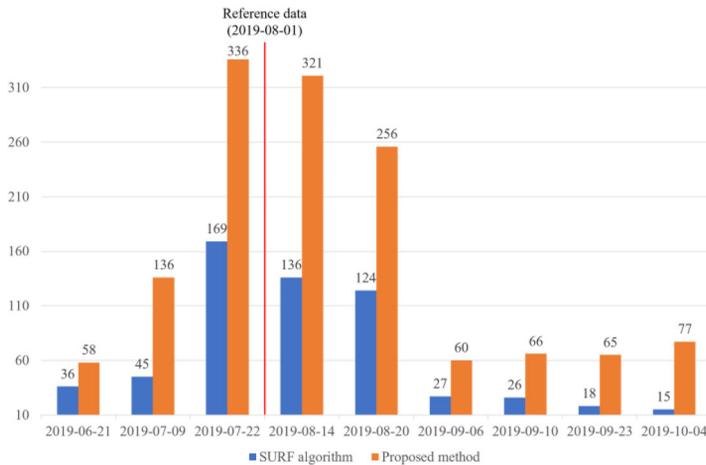


Figure 14. Comparison of extracted number of tie-points by SURF algorithm and proposed method. Note that the numbers are tie-points after removing the mismatched ones; the proposed method’s results are the tie-points extracted during the coarse registration step.

between the images to perform accurate registration. However, the template cannot overlap at similar locations between images when misalignment is severe, such as in the experimental data. For these reasons, inaccurate registration was performed because tie-points were extracted at different locations between the images when using only the MI algorithm. Conversely, the proposed method could extract the tie-point from similar locations between images because it uses the MI algorithm after performing the coarse registration.

The SURF algorithm performed accurate registration on experimental data except for the 2019-09-23 and 2019-10-04 orthophotos, whereas the proposed method obtained the accurate results in all orthophotos. These results are related to the number of extracted tie-points used for performing the image registration. Figure 14 illustrates the trend in the number of tie-points extracted from each method. The number of tie-points extracted from both methods decreased as the difference in acquisition date between the reference and input images increased. Nevertheless, the proposed method extracted approximately two to three times more tie-points than the SURF algorithm in all experimental data. Consequently, the RMSE of coarse registration in the proposed method was 0.382 m that is lower on average than that of the SURF algorithm, and it was particularly estimated to be approximately 1 m lower in 2019-09-23 and 2020-10-04 orthophotos as shown in Table 8.

The proposed DSM registration method corrects the DSM height using EIFs extracted from non-vegetated (height-invariant) areas without additional data. Various experiments

Table 8. Quantitative accuracy calculated by RCP of each method (Unit: m).

Acquisition date	SURF algorithm	SURF with searching space
2019-06-21	0.541	0.312
2019-07-09	0.083	0.071
2019-07-22	0.241	0.023
2019-08-14	0.061	0.055
2019-08-20	0.149	0.081
2019-09-06	0.314	0.227
2019-09-10	0.516	0.372
2019-09-23	1.906	0.485
2019-10-04	1.647	0.386
Average	0.606	0.223

(in Section 4) demonstrated that the proposed method can correct the height of the DSM similarly to the absolute ground height. Therefore, the proposed method has notable versatility and convenience because it only requires accurate orthophotos and DSMs, which are the reference data for correcting the height of DSMs without GCPs.

The corrected DSM using the proposed method can roughly reflect the crop height in the experimental area. The experimental site consists of peppers, soybeans, sesame, perillias, and rice paddies. According to the table of crop characteristics provided by the Rural Development Administration of South Korea, the growth heights of peppers, soybeans, sesame, perillias, and rice paddies average 1.51, 0.75, 1.34, 1.49, and 0.92 m, respectively in Korean standards (Nongsaro 2021). Therefore, the corrected DSM using the proposed method satisfies the height tolerance range (0.75 m–1.51 m) of the experimental site because the maximum RMSEs based on the checkpoints and RCPs are 0.214 and 0.484 m, respectively. However, the height of crops may contain an error since the accuracy of the proposed method is derived from the non-vegetation areas in the experimental site. Nevertheless, the proposed method can effectively correct the DSM without GCPs and be used as a basic method for agricultural applications.

Through the analysis and discussion of various experiments, it has been demonstrated that the performance and accuracy of the proposed method are noteworthy. However, robustness is not guaranteed as these results have been verified using data obtained from a single experimental site. Therefore, we evaluated the robustness by conducting the proposed approach to an additional pair of datasets acquired in Texas A&M AgriLife research centre farm in Corpus Christi, Texas, USA. This dataset was acquired using DJI Phantom 4 Pro with a standard RGB sensor and has 1 inch CMOS 20 Mega Pixel sensor. Reference data was generated using 11 GCPs, while target data was produced using only UAV images. Mosaic images generated with reference and target orthophotos before and after conducting the image registration are given in Figure 15a and b, respectively. In the case of the mosaic image before conducting image registration (Figure 15a), the boundary between each image was expressed inconsistently. On the other hand, the road and crop boundaries in the mosaic images composed of reference and registered orthophotos were aligned shown in Figure 15b. In addition, point clouds of the corrected DSM were converted similarly to the reference DSM as one can see in Figure 15c. The quantitative accuracy of the corrected data was derived as 0.031 m and 0.219 m in the horizontal and vertical directions, respectively (Table 9).

6. Conclusions

In this study, we proposed a method that registers multi-temporal orthophotos and DSMs without GCPs by referring to an accurately produced orthophoto and DSM with GCPs.

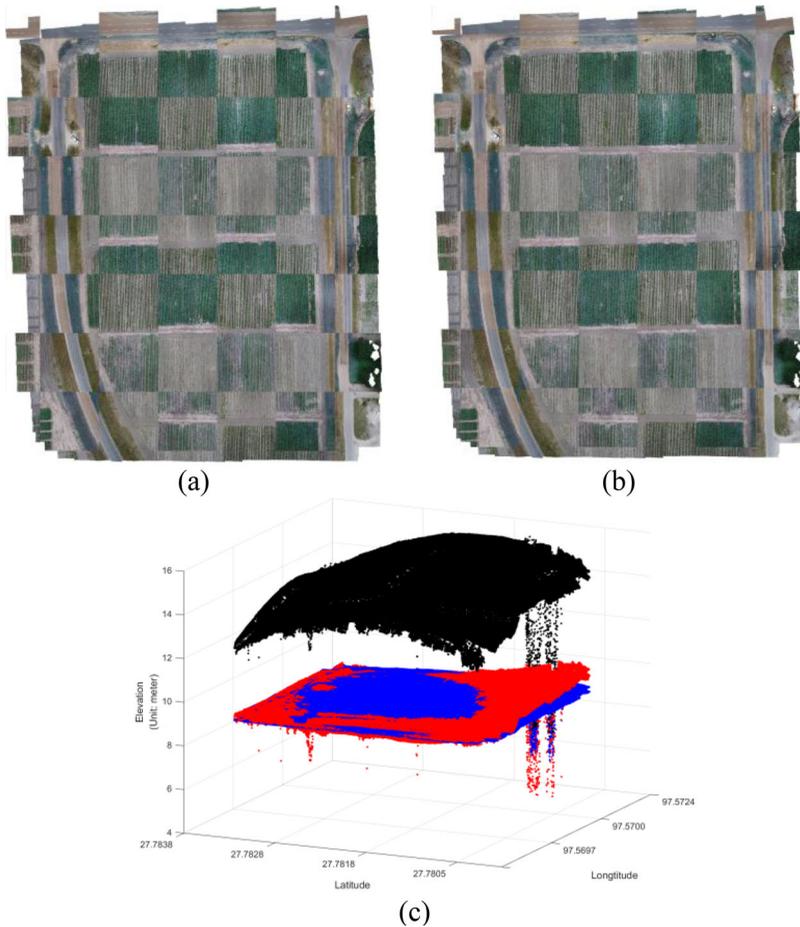


Figure 15. Mosaic images and point clouds experimented on additional dataset: (a) Mosaic image generated with raw orthophotos, (b) Mosaic image generated with registered orthophotos, and (c) Point clouds of DSMs. The reference and target orthophotos are expressed as RGB and BGR in mosaic images, respectively. Point clouds of reference DSM are expressed as red, and point clouds of raw and corrected DSMs are expressed as black and blue, respectively.

Table 9. Determined RMSEs of corrected orthophoto and DSM (Unit: m).

Acquisition date	Horizontal	Vertical
Raw data	4.166	5.410
Registered data	0.031	0.219

The orthophotos produced without GCPs were corrected by performing coarse to fine registration using the SURF and MI algorithms sequentially. Subsequently, the EIFs were extracted from non-vegetation areas by binarizing the ExG image calculated from each orthophoto to register the DSMs. Outliers with weak height invariant properties were removed using the normalized Z-score generated from the height difference between the EIFs for each DSM. Then, the DSM was compensated using a linear regression model based on EIFs. The performance and accuracy of the proposed method were evaluated using multi-temporal orthophotos and DSMs acquired over farmland.

The accuracy of the corrected orthophotos and DSMs using the proposed method was evaluated using the checkpoints obtained from GNSS surveying as assessment data. The average RMSEs based on the checkpoints were 0.024, 0.037, and 0.151 m in the X, Y, and

Z directions, respectively. Furthermore, accurate time-series orthophotos and DSMs were produced using the proposed method. The corrected multi-temporal orthophotos and DSMs had an accuracy of 0.034 m in the horizontal direction and 0.251 m in the vertical direction.

In summary, the proposed method can build high-quality time-series orthophotos and DSMs with suitable geometric accuracy. Therefore, the proposed technique can increase the efficiency of data generation for agricultural applications and has the advantage of using established data. However, because the accuracy of the corrected DSM is an evaluation result based on the height values of ground and road, it is uncertain whether it can effectively reflect the actual crop height. In addition, since the non-vegetation areas in farmland may be dynamically changed, ambiguous errors may thus occur in certain data. Future research will evaluate the applicability and scope of the proposed technique by performing a comparative analysis between the results by monitoring crops using the corrected orthophotos and DSMs and data collected from field observations. Moreover, we will focus on conducting precise DSM registration by identifying the elevation invariant relationship between time-series DSMs.

Author contributions

Conceptualization, T.K. and Y.H.; methodology, T.K. and Y.H.; software, T.K.; validation, T.K.; formal analysis, T.K. and Y.Y.; investigation, T.K., J.P., C.L., and J.J.; data curation, T.K., J.P., C.L., and J.J.; writing—original draft preparation, T.K.; writing—review and editing, Y.H. and J.J.; visualization, T.K. and Y.Y.; supervision, Y.H.; funding acquisition, Y.H. All authors have read and agreed to the published version of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was financially supported by Seoul National University of Science and Technology.

Data availability statement

The data for this study are available upon reasonable request.

References

- Aicardi I, Nex F, Gerke M, Lingua A. 2016. An image-based approach for the co-registration of multi-temporal UAV image datasets. *Remote Sens.* 8(9):779.
- Bastiaanssen WGM, Molden DJ, Makin IW. 2000. Remote sensing for irrigated agriculture: examples from research and possible applications. *Agric WaterManag.* 46(2):137–155.
- Bay H, Ess A, Tuytelaars T, Van Gool L. 2008. Speeded-up robust features (SURF). *Comput Vis Image Underst.* 110(3):346–359.
- Bentoutou Y, Taleb N, Kpalma K, Ronsin J. 2005. An automatic image registration for applications in remote sensing. *IEEE Trans Geosci Remote Sens.* 43(9):2127–2137.
- Chang X, Du S, Li Y, Fang S. 2018. A coarse-to-fine geometric scale-invariant feature transform for large size high resolution satellite image registration. *Sensors.* 18(5):1360.
- Chebrolu N, Labe T, Stachniss C. 2018. Robust long-term registration of UAV images of crop fields for precision agriculture. *IEEE Robot Autom Lett.* 3(4):3097–3104.

- Chen B, Chen Z, Deng L, Duan Y, Zhou J. 2016. Building change detection with RGB-D map generated from UAV images. *Neurocomputing*. 208:350–364.
- Choi HS, Kim EM. 2017. Image registration of drone images through association analysis of linear features. *J Korean Soc Surv Geod Photogramm Cartogr*. 35(6):441–452.
- Du S, Wang M, Fang S. 2017. Block-and-octave constraint SIFT with multi-thread processing for VHR satellite image matching. *Remote Sens Lett*. 8(12):1180–1189.
- Gomez C, Purdie H. 2016. UAV-based photogrammetry and geocomputing for hazards and disaster risk monitoring – a review. *Geoenviron Disasters*. 3(1):1–11.
- Gong M, Zhao S, Jiao L, Tian D, Wang S. 2014. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Trans Geosci Remote Sens*. 52(7):4328–4338.
- Gruszczynski W, Puniach E, Cwiakala P, Matwij W. 2019. Application of convolutional neural networks for low vegetation filtering from data acquired by UAVs. *ISPRS J Photogramm Remote Sens*. 158:1–10.
- Gruszczynski W, Puniach E, Cwiakala P, Matwij W. 2022. Correction of low vegetation impact on UAV-derived point cloud heights with U-Net networks. *IEEE Trans Geosci Remote Sens*. 60:1–18.
- Han Y, Choi J, Byun Y, Kim Y. 2014. Parameter optimization for the extraction of matching points between high-resolution multisensor images in urban areas. *IEEE Trans Geosci Remote Sens*. 52(9):5612–5621.
- Han Y, Choi J, Jung J, Chang A, Oh S, Yeom J. 2019. Automated coregistration of multisensor orthophotos generated from unmanned aerial vehicle platforms. *J Sensors*. 2019:1–10.
- Han Y. 2017. Fine registration between very high resolution satellite images using registration noise distribution. *J Korean Soc Surv Geod Photogramm Cartogr*. 35(3):125–132.
- Hasheminasab SM, Zhou T, Habib A. 2020. GNSS/INS-assisted structure from motion strategies for UAV-based imagery over mechanized agricultural fields. *Remote Sens*. 12(3):351.
- Honkavaara E, Saari H, Kaivosoja J, Pölönen I, Hakala T, Litkey P, Mäkynen J, Pesonen L. 2013. Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture. *Remote Sens*. 5(10):5006–5039.
- Huo C, Pan C, Huo L, Zhou Z. 2012. Multilevel SIFT matching for large-size VHR image registration. *IEEE Geosci Remote Sensing Lett*. 9(2):171–175.
- Kim DW, Yun H, Jeong SJ, Kwon YS, Kim SG, Lee W, Kim HJ. 2018. Modeling and testing of growth status for chinese cabbage and white radish with UAV-based RGB imagery. *Remote Sens*. 10(4):563.
- Kim T, Han Y. 2021. Integrated preprocessing of multitemporal very-high-resolution satellite images via conjugate points-based pseudo-invariant feature extraction. *Remote Sens*. 13(19):3990.
- Kim T, Lee K, Lee WH, Yeom J, Jun S, Han Y. 2019. Coarse to fine image registration of unmanned aerial vehicle images over agricultural area using SURF and mutual information methods. *Korean J Remote Sens*. 35(6–1):945–957.
- Lee HS, Oh JH. 2021. Correcting digital elevation models (DEM) from unmanned aerial vehicles (UAV): a new method using polynomial model matching techniques. *J Coastal Res*. 114(SI):434–438.
- Li J, Hu Q, Ai M, Zhong R. 2017. Robust feature matching via support-line voting and affine-invariant ratios. *ISPRS J Photogramm Remote Sens*. 132:61–76.
- Li M, Shamshiri RR, Schirrmann M, Weltzien C, Shafian S, Laursen MS. 2022. UAV oblique imagery with an adaptive micro-terrain model for estimation of leaf area index and height of maize canopy from 3D point clouds. *Remote Sens*. 14(3):585.
- Liu J, Xu W, Guo B, Zhou G, Zhu H. 2021. Accurate mapping method for UAV photogrammetry without ground control points in the map projection frame. *IEEE Trans Geosci Remote Sens*. 59(11):9673–9681.
- Maes WH, Steppe K. 2019. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci*. 24(2):152–164.
- Mulla DJ. 2013. Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. *Biosyst Eng*. 114(4):358–371.
- Na S, Park C, So K, Ahn H, Lee K. 2018. Application method of unmanned aerial vehicle for crop monitoring in Korea. *Korean J Remote Sens*. 34(5):829–846.
- Nex F, Remondino F. 2014. UAV for 3D mapping applications: a review. *Appl Geomat*. 6(1):1–15.
- Nongsaro. 2021. Information of crops varieties. 2021. Repub Korea Rural Dev Adm; [accessed 2020 May 27]. <http://www.nongsaro.go.kr>.
- Oh J, Han Y. 2020. A double epipolar resampling approach to reliable conjugate point extraction for accurate kompsat-3/3A stereo data processing. *Remote Sens*. 12(18):2940.
- Paul S, Pati UC. 2021. A comprehensive review on remote sensing image registration. *Int J Remote Sens*. 42(14):5396–5432.

- Raeva PL, Šedina J, Dlesk A. 2019. Monitoring of crop fields using multispectral and thermal imagery from UAV. *Eur J Remote Sens.* 52(sup1):192–201.
- Rokhmana CA. 2015. The potential of UAV-based remote sensing for supporting precision agriculture in indonesia. *Procedia Environ Sci.* 24:245–253.
- Torres-Sánchez J, Peña JM, de Castro AI, López-Granados F. 2014. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Comput Electron Agric.* 103:104–113.
- Tsai CH, Lin YC. 2017. An accelerated image matching technique for UAV orthoimage registration. *ISPRS J Photogramm Remote Sens.* 128:130–145.
- Tsouros DC, Bibi S, Sarigiannidis PG. 2019. A review on UAV-based applications for precision agriculture. *Inf.* 10(11):349.
- Uysal M, Toprak AS, Polat N. 2015. DEM generation with UAV photogrammetry and accuracy analysis in sahitler hill. *Meas.* 73:539–543.
- Viola P, Wells WM. 1997. Alignment by maximization of mutual information. *Int J Comput Vis.* 24(2):137–154.
- Wang C, Myint SW. 2015. A simplified empirical line method of radiometric calibration for small unmanned aircraft systems-based remote sensing. *IEEE J Sel Top Appl Earth Observ Remote Sens.* 8(5):1876–1885.
- Wang C, Wang L, Liu L. 2014. Progressive mode-seeking on graphs for sparse feature matching. In: *Proceeding of the European Conference on Computer Vision; Sep 6–12; Zurich.* p.788–802.
- Wei Z, Han Y, Li M, Yang K, Yang Y, Luo Y, Ong SH. 2017. A small UAV based multi-temporal image registration for dynamic agricultural terrace monitoring. *Remote Sens.* 9(9):904.
- Woebbecke DM, Meyer GE, Von Bargen K, Mortensen DA. 1995. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans Am Soc Agric Eng.* 38(1):259–269.
- Xiang H, Tian L. 2011. Method for automatic georeferencing aerial remote sensing (RS) images from an unmanned aerial vehicle (UAV) platform. *Biosyst Eng.* 108(2):104–113.
- Ye Z, Kang J, Yao J, Song W, Liu S, Luo X, Xu Y, Tong X. 2020. Robust fine registration of multisensor remote sensing images based on enhanced subpixel phase correlation. *Sensors.* 20(15):4338.
- Yeom J, Jung J, Chang A, Ashapure A, Maeda M, Maeda A, Landivar J. 2019. Comparison of vegetation indices derived from UAV data for differentiation of tillage effects in agriculture. *Remote Sens.* 11(13):1548.
- Yun H. 2017. Development of remote sensing technology for growth estimation of white radish and napa cabbage using UAV and RGB camera. [master's thesis]. Seoul: Seoul National University.
- Zhang C, Kovacs JM. 2012. The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agric.* 13(6):693–712.
- Zhang T, Zhao R, Chen Z. 2020. Application of migration image registration algorithm based on improved SURF in remote sensing image mosaic. *IEEE Access.* 8:163637–163645.
- Zhang Y, Xia C, Zhang X, Cheng X, Feng G, Wang Y, Gao Q. 2021. Estimating the maize biomass by crop height and narrowband vegetation indices derived from UAV-based hyperspectral images. *Ecol Indic.* 129:107985.
- Zhuo X, Koch T, Kurz F, Fraundorfer F, Reinartz P. 2017. Automatic UAV image geo-registration by matching UAV images to georeferenced image data. *Remote Sens.* 9(4):376.
- Zitová B, Flusser J. 2003. Image registration methods: a survey. *Image Vis Comput.* 21(11):977–1000.