

Article

Robust Mosaicking of Lightweight UAV Images Using Hybrid Image Transformation Modeling

Jae-In Kim ^{1,2} , Hyun-cheol Kim ¹  and Taejung Kim ^{2,*}

¹ Unit of Arctic Sea-Ice Prediction, Korea Polar Research Institute (KOPRI), Incheon 21990, Korea; jikim0312@kopri.re.kr (J.-I.K.); kimhc@kopri.re.kr (H.-c.K.)

² Department of Geoinformatic Engineering, Inha University, Incheon 21990, Korea

* Correspondence: tejid@inha.ac.kr; Tel.: +82-32-860-7606

Received: 23 January 2020; Accepted: 18 March 2020; Published: 20 March 2020



Abstract: This paper proposes a robust feature-based mosaicking method that can handle images obtained by lightweight unmanned aerial vehicles (UAVs). The imaging geometry of small UAVs can be characterized by unstable flight attitudes and low flight altitudes. These can reduce mosaicking performance by causing insufficient overlaps, tilted images, and biased tiepoint distributions. To solve these problems in the mosaicking process, we introduce the tiepoint area ratio (TAR) as a geometric stability indicator and orthogonality as an image deformation indicator. The proposed method estimates pairwise transformations with optimal transformation models derived by geometric stability analysis between adjacent images. It then estimates global transformations from optimal pairwise transformations that maximize geometric stability between adjacent images and minimize mosaic deformation. The valid criterion for the TAR in selecting an optimal transformation model was found to be about 0.3 from experiments with two independent image datasets. The results of a performance evaluation showed that the problems caused by the imaging geometry characteristics of small UAVs could actually occur in image datasets and showed that the proposed method could reliably produce image mosaics for image datasets obtained in both general and extreme imaging environments.

Keywords: lightweight UAV; image mosaic; imaging geometry; tiepoint area ratio

1. Introduction

Lightweight unmanned aerial vehicles (UAVs) are widely used as a remote sensing platform for obtaining high spatial resolution images. Low-altitude flight and easy control are the most distinctive features of UAVs, compared to conventional remote sensing platforms, such as aircrafts and satellites. On the other hand, the improvement in spatial resolution causes a reduction in the ground area that can be covered by a single image. This means that many UAV images are needed in order to analyze wide target areas. As a result, image mosaicking is regarded as an essential task in UAV applications.

Image mosaicking methods can be classified into spatial-data-based methods and feature-based methods. The former methods generate mosaic images using digital surface models (DSMs) or ground control points (GCPs) [1–6]. The latter methods generate mosaic images based on tiepoints between adjacent images [7–10]. In many remote sensing applications, spatial-data-based methods are preferred because they can produce georeferenced or ortho-rectified mosaic images. However, feature-based methods can also be used effectively in investigating disaster regions, such as fire, flood, and earthquake zones; and polar regions, such as icebergs, glaciers, and sea-ice. In these cases, it is of paramount importance to quickly report on-site situations to decision-makers. To do this, spatial-data-based methods require excessive time for constructing DSMs or GCPs [11,12].

In this paper, we present a feature-based mosaicking method to further improve the utilization of lightweight UAVs in extreme environments. Existing studies made an effort to improve the

accuracy and speed of image mosaicking. Enhancing mosaicking speed was tried by reducing the processing time for tiepoint extraction, which requires the greatest amount of computation. Moussa and El-Sheimy [13] minimized the number of matching image pairs by structuring UAV images through Delaunay triangulation. Mehrdad et al. [14] reduced tiepoint extraction regions using epipolar geometry established from initial camera parameters. Faraji et al. [15] minimized the computation of tiepoint extraction using reference images in target areas. On the other hand, enhancing mosaicking accuracy was tried by minimizing the accumulated errors and distortions that are likely to occur in image transformation estimation. Moussa and El-Sheimy [13] minimized the accumulated errors using proximity among the images. Xu et al. [9] minimized image distortions by making image transformations as close as possible to rigid transformations. Mehrdad et al. [14] mitigated accumulated errors by estimating image transformations so that reprojection errors of tiepoints are minimized.

The studies described above contributed to improving the performance of image mosaicking. However, considerations for poor imaging environments were not fully discussed. The existing methods assumed high overlaps and well-distributed tiepoints between adjacent images to establish image transformations. The conventional remote sensing platforms can easily meet these requirements, but lightweight UAVs might not. Because UAVs are sensitive to changes in wind direction and speed, overlaps may not be well maintained between adjacent images taken during a flight [16,17]. In addition, since UAVs have low flight altitudes, tiepoint distributions may be biased, especially with low-textured surfaces [16]. In this situation, even if there are sufficient overlaps, the accuracy of transformations would be reduced due to the biased tiepoint distributions.

Therefore, in this paper, we investigate a new, robust mosaicking method that can handle problems caused by the imaging geometry characteristics of lightweight UAVs. In one of our previous studies, we found that applying a simple transformation model, such as an affine model, could yield better results for narrowly overlapping image pairs than by applying a sophisticated model, such as a homography model [18]. In the subsequent study, we examined the possibility for the selective use of transformation models [19]. Based on these findings, we propose an image mosaicking method that can establish optimal transformations and also minimize mosaic deformation. In addition, we experimentally demonstrate the problems and analyze the effects on mosaicking results. The proposed method estimates pairwise transformations between adjacent images. The optimal transformation model for each pair is derived from a geometric stability indicator that can consider both overlap and tiepoint distribution simultaneously. The proposed method then estimates global transformations from optimal pairwise transformations that maximize geometric stability between adjacent images and minimize mosaic deformation. The criterion for assessing geometric stability in selecting an optimal transformation model was determined through experiments using two independent image datasets. Performance evaluations were conducted using a highly overlapping image dataset and an inconsistently overlapping image dataset.

2. Materials and Methods

2.1. Dataset

Three image datasets were used to develop and evaluate the proposed method: Dataset-1A, Dataset-1B, and Dataset-2. These datasets were obtained by experts with flight authorizations for research purposes. Dataset-1A and Dataset-1B consist of single strip images, which were acquired by a small drone, a DJI S900 (DJI, Shenzhen, China), with a total weight of 3.3 kg and a maximum flight time of 18 minutes. The images in Dataset-1A and Dataset-1B were used to determine the criterion of geometric stability for selecting optimal transformation models between adjacent images. Figure 1 and Table 1 show the image acquisition information for the two datasets.



Figure 1. Flight lines and exposure points for the two image strips: (a) the flight plan for Dataset-1A and (b) the flight plan for Dataset-1B. These figures were formulated by editing Figure 5; Figure 15 from the work of Kim et al. [18].

Table 1. Detailed specifications for Dataset-1A and Dataset-1B.

Name	Location	No. of Images	Image Sensor	Focal Length	Pixel Size	Image Size	Overlap	Flight Height
Dataset-1A	Incheon, Korea	11	Sony EOS 60D	40 mm	4.3 μm	5184 \times 3456	64–80%	250 m
Dataset-1B	Daegu, Korea	7	Sony ILCE-7R	35 mm	4.9 μm	4800 \times 3200	63–78%	187 m

Dataset-2 consists of multistrip images, which were obtained by a small drone, SmartOne (SmartPlanes, Skellefteå, Sweden), with a total weight of 1.5 kg and a maximum flight time of 50 min. The images in Dataset-2 were used to evaluate the proposed method. Figure 2 and Table 2 show the detailed information for Dataset-2. Performance evaluations were conducted for both general and poor imaging environments. For a general imaging environment, all images in Dataset-2 were used, as seen in Figure 2a. For an extreme imaging environment, a subset of Dataset-2 with inconsistent overlaps was used, as seen in Figure 2b.

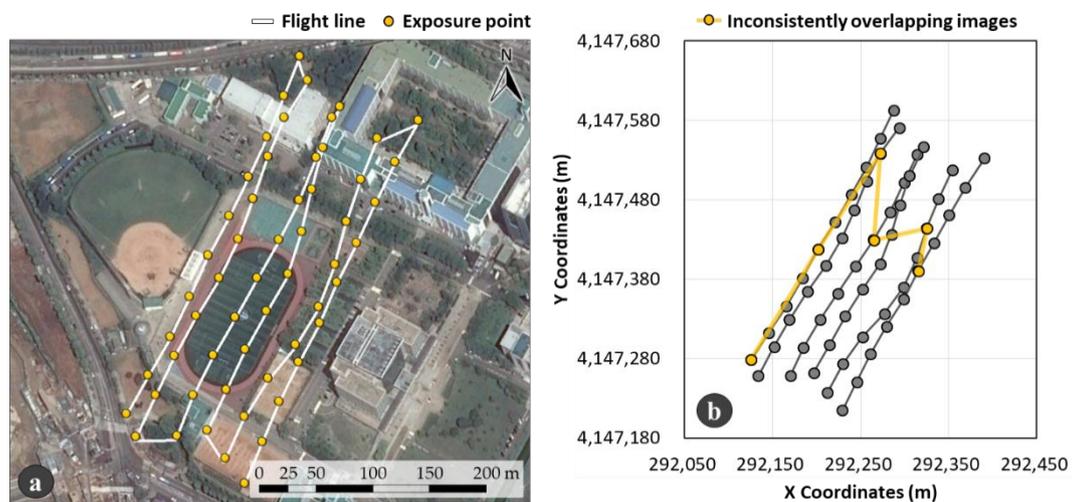


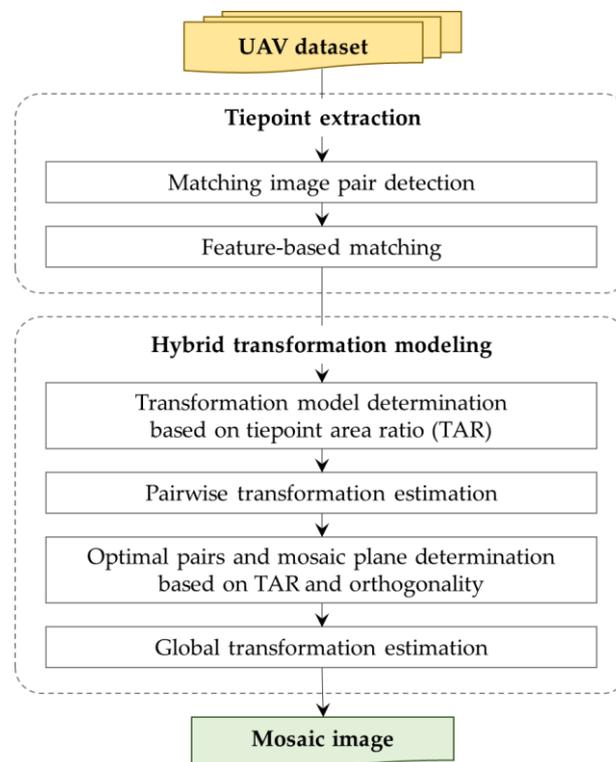
Figure 2. Flight lines and exposure points for Dataset-2: (a) the flight plan for Dataset-2 and (b) inconsistently overlapping image subset.

Table 2. Detailed specifications for Dataset-2.

Name.	Target Area	Strips/Images	Sensor	Focal Length	Pixel Size	Image Size	Overlaps	Flight Height
Dataset-2	Incheon, Korea	6/57	Ricoh GR II	18 mm	4.8 μ m	4928 \times 3264	64–96%	158 m

2.2. Proposed Method

The proposed mosaicking method consists of two parts: tiepoint extraction and hybrid transformation modeling. Conventional mosaicking methods often include an image-blending process to mitigate the discrepancies between pixel values on overlapping image regions. However, because we focus on robustness to poor imaging environments, this additional process is not considered here. Figure 3 shows the workflow of the proposed method. The whole process for image mosaicking and evaluations was implemented in the C++ language and OpenCV library (ver. 2.4.9). For statistics and graph visualization, Microsoft Office 2016 was used.

**Figure 3.** Workflow of the proposed mosaicking method.

In the first part of our method, tiepoints were extracted to estimate pairwise transformations between adjacent image pairs. This process starts with determining matching image pairs to extract tiepoints. This is to avoid unnecessary computations on nonoverlapping image pairs. To achieve this, exterior orientation parameter (EOP)-based methods [14,20], the Delaunay triangulation-based method [13], and a graph-based method [21] can be used. From among them, we employed the EOP-based method using data obtained by a global positioning system/inertial navigation system mounted on a UAV. After determining matching pairs, tiepoints were extracted using a feature-based method. Because tiepoint extraction for UAV images has to achieve not only quickness in processing a large number of images but also robustness to changes in rotation and scale between images, we adopted the fast retina keypoint (FREAK) method, which is known among binary descriptor methods to be invariant to rotation and scale changes [22].

After tiepoint extraction, image transformations were established via hybrid transformation modeling. We divided image transformations into pairwise (image-to-image) transformations and global (image-to-mosaic) transformations. Because global transformations are derived from pairwise transformations, the performance from image mosaicking depends on the accuracy of pairwise transformations [7–9]. In our proposed method, a pairwise transformation can be established between two transformation models: affine transformation and homography models.

Affine transformation with six degrees of freedom (DOF) can describe scale, rotation, translation, and skew between two image planes in two-dimensional (2D) space as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

where (x, y) and (x', y') are the image coordinates and their transformed coordinates, respectively. On the other hand, homography with eight DOF can explain general motions between two image planes in 3D space as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

Consequently, the homography model is generally known to be more appropriate than the affine transformation model in estimation of pairwise transformations [7,9]. However, this presupposes high overlaps and well-distributed tiepoints between adjacent images, whereas small UAVs may not meet those requirements due to unstable flight attitudes and low flight altitudes. In these cases, conventional methods for estimating pairwise transformation may make it difficult to produce reliable results. Figure 4 is an example showing that the distribution of tiepoints can be biased due to a decrease in flight altitude. We can see that the features in the aircraft image are evenly distributed (Figure 4a), while the features in the UAV image are concentrated in some areas (Figure 4b). Features can usually be extracted from textured surfaces in images, and tiepoints between adjacent images are determined by matching features of each image. Therefore, if textures are nonuniformly distributed for overlapping areas, distributions of tiepoints may also be biased. In this regard, the aircraft image would have well-distributed tiepoints for any overlapping areas if it has sufficient overlaps with other images. On the other hand, the UAV image may have biased tiepoint distributions if overlapping areas are formed on the left side with low-textured surfaces. This suggests that UAV images may have a relatively high proportion of low-textured surfaces for overlapping areas due to low flight altitudes, and thus, distributions of tiepoints may also be more biased. Largely biased tiepoint distributions would cause transformations over-fitted for some areas.

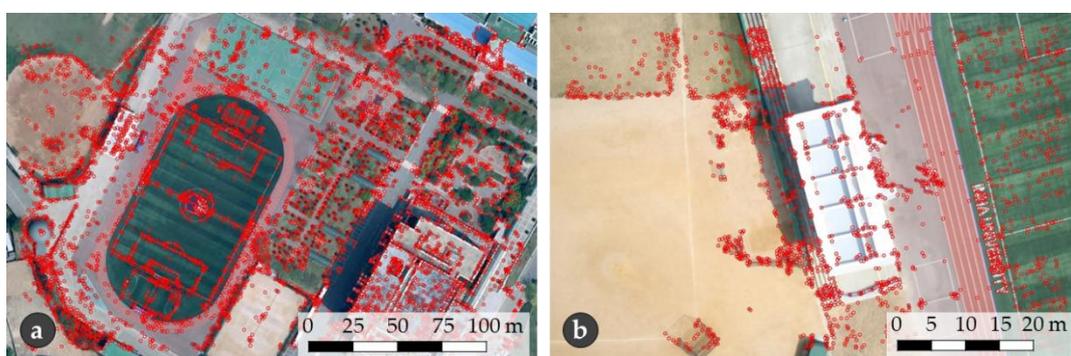


Figure 4. Distribution of feature points (red circles) extracted from (a) an image taken from an aircraft and (b) an image taken from a UAV.

For these reasons, in our method, pairwise transformation between two images is established with an optimal model selected by geometric stability analysis. For high geometric stability, homography is applied as a precision model, and for low geometric stability, affine transformation is applied as a robust model. As a confidence indicator for geometric stability analysis, overlapping area ratio (OAR) or number of tiepoints (NoT) can be used. However, these indicators cannot represent tiepoint distributions. Thus, we introduced a new geometric stability indicator, tiepoint area ratio (TAR), which can simultaneously consider both overlap and tiepoint distribution. This indicator is defined as the ratio of the tiepoint area to the entire image area, as shown in Figure 5. The tiepoint area means overlapping regions affected by tiepoints and consists of Delaunay triangles formed by tiepoints. Thus, the TAR is formulated as

$$\text{TAR} = \frac{\sum_{i=1}^N |(x_2^i - x_1^i)(y_3^i - y_1^i) - (x_3^i - x_1^i)(y_2^i - y_1^i)|}{WH} \quad (3)$$

where W and H are the width and height of the image, respectively, N is the number of Delaunay triangles, and (x_1^i, y_1^i) , (x_2^i, y_2^i) , and (x_3^i, y_3^i) are the image coordinates for the three vertices of the i th triangle. The criterion for the TAR in selecting an optimal transformation model was determined by correlation analysis between transformation errors and the TAR values. The experiments are covered in Section 3.1.

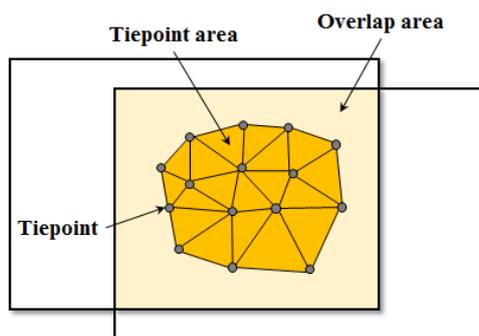


Figure 5. Conceptual diagram for tiepoint area ratio (TAR).

Global transformations of individual images can be established by concatenating pairwise transformations between adjacent images. Because this process may cause error propagation in pairwise transformations, an optimization method is required to minimize error accumulation. To this end, graph methods [23–25] and bundle adjustment methods [7,9,15,26,27] have been proposed. We adopted a modified graph method to ensure efficiency and robustness in image mosaicking. Graph methods generally consist of maximum spanning tree (MST) generation and mosaic plane selection. An MST indicates optimal image pairs that minimize error accumulation in concatenating pairwise transformations. This is generally derived from NoT as a weight [21,28]. A mosaic plane means a common 2D plane to reproject raw images. This is determined by a root image that minimizes the depth of the MST derived [23–25]. In general cases, graph methods would lead to satisfactory mosaicking results. However, in extreme imaging environments, they may not guarantee acceptable performance. Thus, we modified an existing graph method to consider the imaging geometry characteristics of small UAVs. For MST generation, the TAR was applied as a weight instead of NoT or OAR. This aims to prevent situations where unreliable pairwise transformations are involved in estimating global transformations. For mosaic plane selection, an image that minimizes the deformations of reprojected images was used as a mosaic plane. This avoids unnecessary mosaic

deformation that can occur by selecting a largely tilted image as a mosaic plane. Mosaic deformation is calculated from the orthogonality of reprojected images as follows:

$$\text{Mosaic deformation} = \sqrt{\frac{\sum_{i=1}^N (\theta'_i - 90^\circ)^2}{N}}, \quad \theta' = \cos^{-1} \left(\frac{\vec{X}' \cdot \vec{Y}'}{\|\vec{X}'\| \|\vec{Y}'\|} \right) \tag{4}$$

where N is the number of images and θ'_i indicates the orthogonality of the transformed i th image. Orthogonality means the angle between the two image axes, as seen in Figure 6 [29]. Figure 7a,b shows the mosaicking results when the mosaic plane is properly selected and when it is not, respectively.

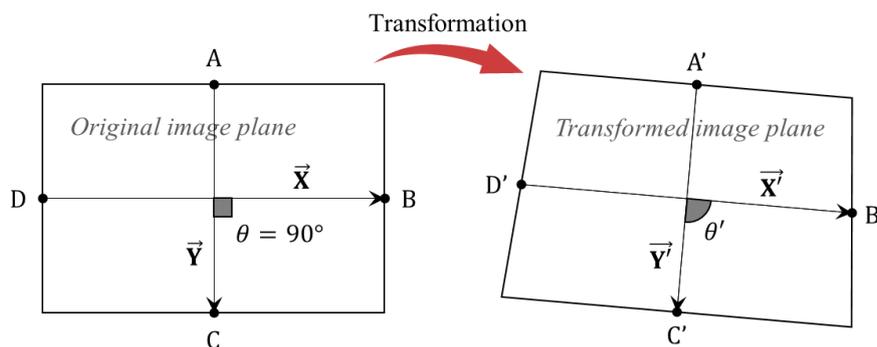


Figure 6. Orthogonality for the original image and transformed image.

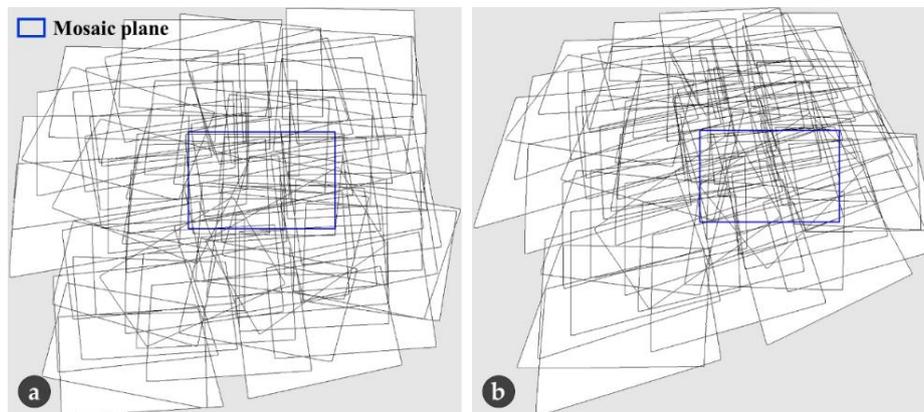


Figure 7. Differences in mosaics by mosaic plane selection: (a) mosaic by selecting an appropriate image as a mosaic plane and (b) mosaic by selecting a largely tilted image as a mosaic plane.

2.3. Evaluation indicators

The proposed method was evaluated from mosaicking errors and distortions. Mosaicking errors are measured for pairwise and global transformations. Because the proposed method generates image mosaics using only tiepoints between adjacent images, mosaicking errors to evaluate geometric performance are calculated from reprojection errors between adjacent images [21,24,25]. The reprojection error is defined by the Euclidean distance between observed tiepoints and calculated tiepoints, as follows:

$$\text{Reprojection error} = \frac{\sum_{n=1}^N \sqrt{(x_n - \hat{x}_n)^2 + (y_n - \hat{y}_n)^2}}{N} \tag{5}$$

where N is the total number of tiepoints and (x_n, y_n) and (\hat{x}_n, \hat{y}_n) are the observed and calculated image coordinates for the n th tiepoint, respectively.

Mosaicking distortions are measured from orthogonality differences with reference images derived by camera parameters of images. The camera parameters, including interior (focal length, pixel size, principal points, and lens distortion coefficients) and exterior parameters (positions and orientations), were obtained by commercial software, Pix4D (ver. 4.4.12). Figure 8 shows the mosaic generated by the reference images. The red lines indicate the boundaries of the individual reference images.

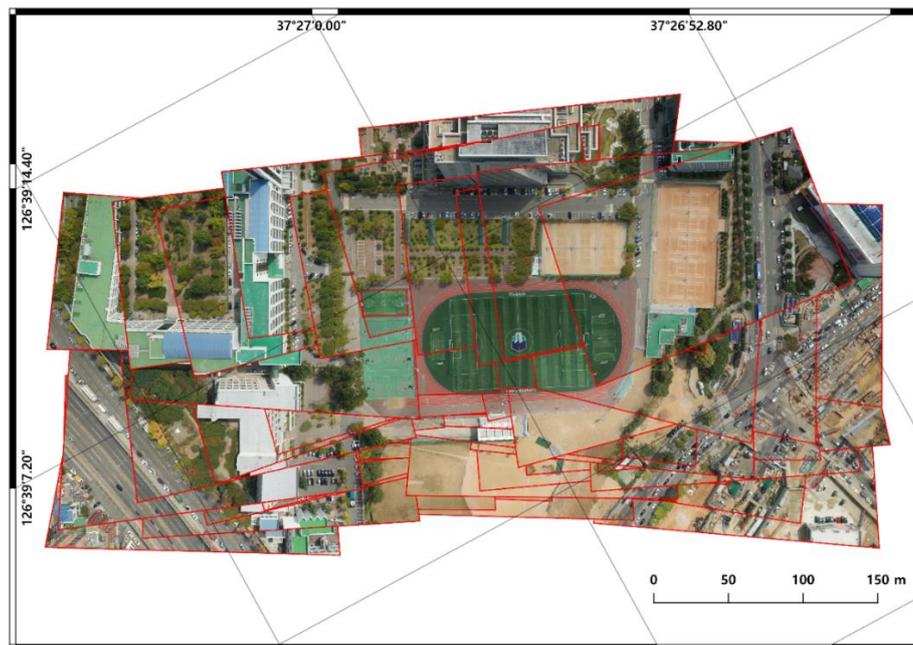


Figure 8. Reference images mosaicked by the interior and exterior camera parameters.

3. Results and Discussion

3.1. Criterion Determination for Hybrid Transformation Modeling

The criterion for the TAR in selecting an optimal transformation model was determined by correlation analysis using two independent single image strips in Dataset-1A and Dataset-1B. For correlation analysis between errors in pairwise transformations and values of TAR, many image pairs with different overlaps and tiepoint distributions are required. To achieve this, we created additional image pairs with different conditions from the raw image datasets through overlap adjustment. Overlap adjustment was performed by removing the outer parts of images to preserve the perspective property of the frame images. As a result, 710 image pairs and 768 image pairs were produced from Dataset-1A and Dataset-1B, respectively. Model tiepoints for estimation of pairwise transformations were automatically extracted from raw image pairs using the FREAK algorithm. Check tiepoints for evaluation of transformations were manually obtained. Table 3 shows the number of tiepoints extracted from raw image pairs. Figure 9 illustrates the number of tiepoints used for estimating pairwise transformations of overlap-adjusted image pairs. Transformations for overlap-adjusted image pairs were estimated using affine transformation and homography models from model tiepoints within their overlapping areas. On the other hand, transformations were evaluated for all check tiepoints, not only in actually overlapping areas but also in truncated areas. This was intended to analyze transformation errors consistently for all overlap-adjusted image pairs. Transformation errors were

measured as reprojection errors, defined by the Euclidean distance between observed tiepoints and calculated tiepoints, as follows:

$$\text{Reprojection error} = \frac{\sum_{n=1}^N \sqrt{(x_n - \hat{x}_n)^2 + (y_n - \hat{y}_n)^2}}{N} \tag{6}$$

where N is the total number of check tiepoints and (x_n, y_n) and (\hat{x}_n, \hat{y}_n) are the observed and calculated image coordinates for the n th check tiepoint, respectively.

Table 3. Model and check tiepoints extracted from the original image pairs.

Number of Tiepoints	Model Tiepoints		Check Tiepoints	
	Dataset-1A	Dataset-1B	Dataset-1A	Dataset-1B
Minimum	108	142	187	169
Mean	137	187	327	235
Maximum	215	259	459	270

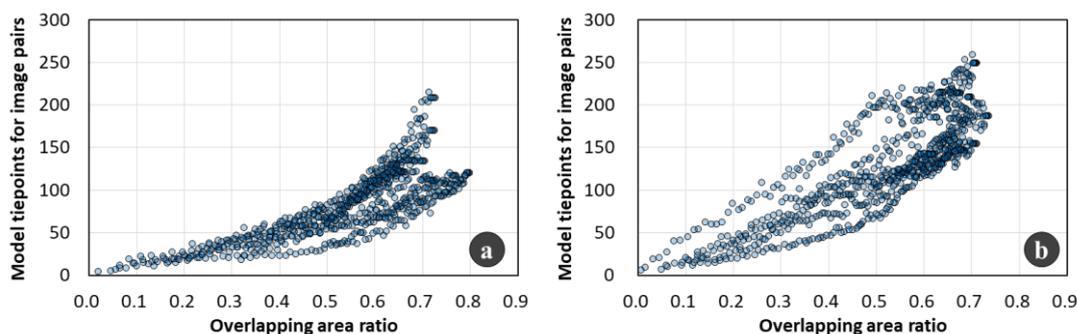


Figure 9. Model tiepoints for overlap-adjusted image pairs: (a) Dataset-1A and (b) Dataset-1B.

The results of correlation analysis between errors of transformation and values of geometric stability indicators, such as NoT, OAR, and TAR, are summarized in Table 4. In these results, reprojection errors increased rapidly as the values of geometric stability indicators decreased, so their relationships could be modeled as power function forms. So far, many studies have used NoT or OAR to evaluate the geometric stability between adjacent images [21,23,28–30]. However, as seen in Figure 10, NoT and OAR showed relatively large uncertainties in the low geometric stability range (i.e., adjacent images with small number of tiepoints or narrow overlaps). These results indicate that NoT and OAR may not be able to reliably evaluate the geometric stability, especially between UAV images. On the other hand, TAR, which can simultaneously consider both overlap and tiepoint distribution, showed the highest correlation with reprojection errors. In addition, the TAR appropriately reflected changes in reprojection errors, even in the low geometric stability range. These results demonstrate that TAR can be used effectively as a geometric stability indicator in estimating pairwise transformations of UAV images. Consequently, the criterion for the TAR could be determined by comparing two regression models for affine transformation and homography. As seen in Figure 11, homography-based transformations had smaller reprojection errors in the high TAR range than those from affine transformation, whereas affine-based transformations had smaller reprojection errors in the low TAR range. The TAR value at the reversal point of reprojection errors was about 0.3, and the results were found to be the same for both Dataset-1A and Dataset-1B. Therefore, based on these results, we determined that a TAR value of 0.3 is a reliable criterion for selecting an optimal transformation model.

Table 4. Correlation analysis results from the number of tiepoints (NoT), overlapping area ratio (OAR), and tiepoint area ratio (TAR) for Dataset-1A and Dataset-1B.

	Models	NoT	OAR	TAR
Dataset-1A	Affine Transformation	$y = 80.20x^{-0.56}$ $R^2 = 0.56$ Rography – LMOrmation	$y = 4.49x^{-0.72}$ $R^2 = 0.46$ Rography – LMOrmation	$y = 3.81x^{-0.55}$ $R^2 = 0.60$ Rography – LMOrmation
	Homography	$y = 193.14x^{-0.77}$ $R^2 = 0.56$ Rography – LMOrmation	$y = 2.89x^{-1.34}$ $R^2 = 0.63$ Rography – LMOrmation	$y = 2.63x^{-0.88}$ $R^2 = 0.76$ Rography – LMOrmation
Dataset-1B	Affine Transformation	$y = 53.58x^{-0.36}$ $R^2 = 0.22$ Rography – LMOrmation	$y = 6.91x^{-0.48}$ $R^2 = 0.23$ Rography – LMOrmation	$y = 4.87x^{-0.48}$ $R^2 = 0.43$ Rography – LMOrmation
	Homography	$y = 88.10x^{-0.47}$ $R^2 = 0.19$ Rography – LMOrmation	$y = 5.73x^{-0.72}$ $R^2 = 0.20$ Rography – LMOrmation	$y = 3.44x^{-0.77}$ $R^2 = 0.47$ Rography – LMOrmation

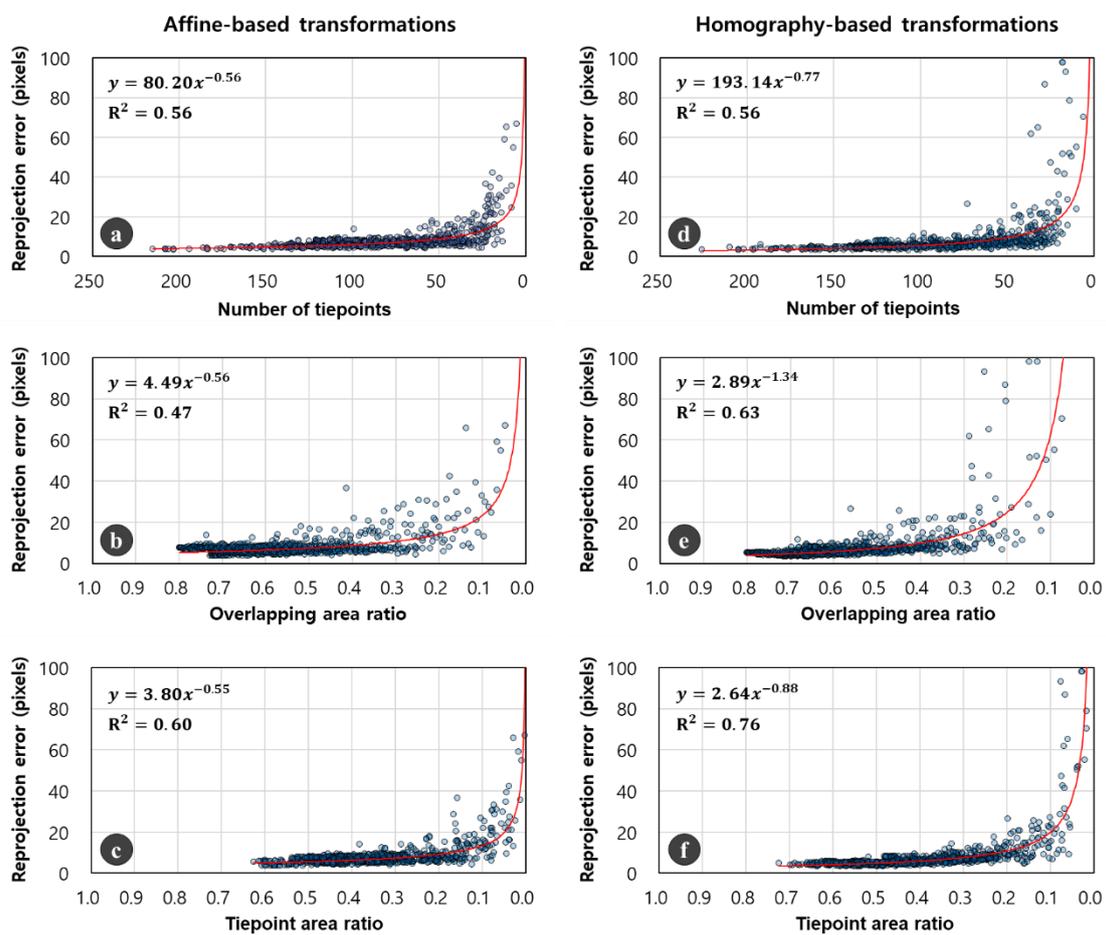


Figure 10. Scatterplots between errors of transformation and values of geometric stability indicators for Dataset-1A: (a,d) scatterplots for the number of tiepoints; (b,e) scatterplots for overlapping area ratio; and (c,f) scatterplots for tiepoint area ratio. Pairwise transformations were estimated using (a–c) the affine transformation model and (d–f) the homography model.

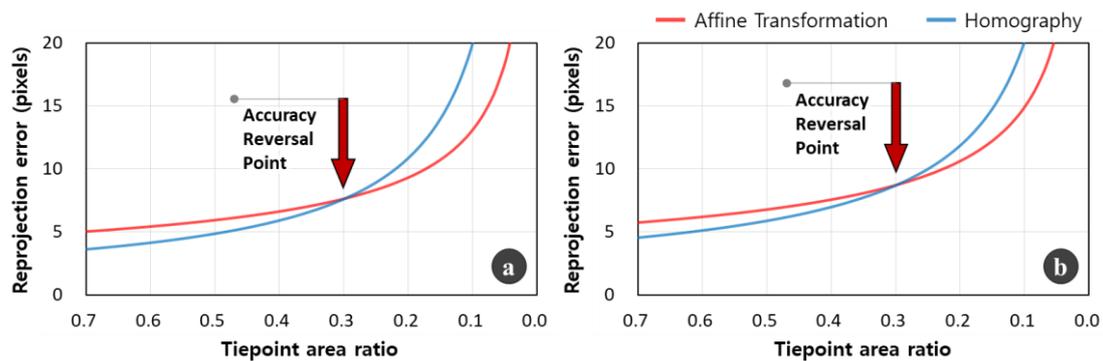


Figure 11. Comparison of the two regression models derived by affine transformation and homography: (a) Dataset-1A and (b) Dataset-1B.

3.2. Mosaicking Performance Evaluation

We analyzed mosaicking performance for both general and poor imaging environments, as seen in Figure 2. In addition, we compared the proposed method with traditional affine transformation-based and homography-based mosaicking methods. In these comparative methods, MSTs were generated by NoT and OAR, respectively, and mosaic planes were determined by root images that minimize the depth of the MSTs. These comparative methods were also implemented by ourselves in the C++ language and OpenCV library (ver. 2.4.9).

Model tiepoints for estimation of pairwise transformations were acquired using the FREAK algorithm, as explained in the experiment described above. On the other hand, check tiepoints for performance evaluation were extracted with two-step processing. We first extracted initial tiepoints using the scale invariant feature transform (SIFT) algorithm [31], which takes more processing time but allows more accurate tiepoint extraction. We then selected multiple tiepoints observed in three or more images as checkpoints. Although this extraction procedure cannot obtain a large number of tiepoints, it secures reliable tiepoints. Table 5 shows the results of tiepoint extraction.

Table 5. The model and check tiepoints for Dataset-2.

	Number of Tiepoints for Image Pairs				Number of Image Pairs
	Minimum	Mean	Maximum	Total	
Model points	7	208	1486	132,054	684
Check points	10	24	77	6101	251

The evaluation results for the general imaging environment are summarized in Table 6, where the ranges of NoT, OAR, and TAR were calculated for optimal image pairs, and the mosaicking errors were calculated for all adjacent image pairs. The reprojection errors in the evaluation results may appear relatively larger than the results of existing studies [21,25]. However, these are due to large relief displacements by high elevation changes and low flight altitudes. Note that the existing studies mostly used images taken at high altitudes for flat areas.

Table 6. Mosaicking performance evaluations for the general imaging environment.

Pairwise Modeling	Global Modeling		NoT			OAR			TAR			Reprojection Errors (Pixels)		Distort. (deg.)
	MST Weight	Base Plane	Min.	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.	Pairwise	Global	
Affine	NoT	NoT	187	742	1486	0.54	0.72	0.96	0.32	0.56	0.85	51.80	475.60	6.70
Homo.	NoT	NoT	187	742	1486	0.50	0.72	0.96	0.32	0.56	0.85	15.73	59.88	5.08
Affine	OAR	OAR	6	543	1486	0.66	0.78	0.98	0.07	0.55	0.85	51.80	191.85	5.34
Homo.	OAR	OAR	34	545	1486	0.64	0.78	0.96	0.26	0.55	0.85	15.73	35.16	6.06
Hybrid	TAR	Ortho	143	661	1486	0.61	0.76	0.96	0.38	0.60	0.85	42.68	18.78	4.01

In this experiment, the homography model was more effective than the affine transformation model for pairwise transformation modeling, as reported in existing studies [7,9], and OAR was more appropriate than NoT for MST generation and mosaic plane selection. Consequently, from among the comparative methods, the mosaicking method using the homography model and OAR showed the best performance.

Meanwhile, the proposed method showed about two times better performance than the best of the comparative methods. In this experiment, the proposed method had to establish all pairwise transformations of optimal image pairs from the homography model, because all TAR values derived for the optimal image pairs were higher than the criterion for hybrid transformation modeling (i.e., a TAR value of 0.3). This means that the performance enhancement from the proposed method was caused by global transformation modeling. Thus, we could know that mosaicking accuracy can vary greatly, depending on how to construct the optimal image pairs. This implies that if there are image pairs with large errors between the optimal image pairs derived, they will greatly propagate the errors to all the images connected to them [32]. In fact, all cases with the same transformation model yielded the same pairwise errors while producing different global errors. Note that the relatively large pairwise error of the proposed method was due to the affine-based transformations that were excluded in MST generation. Therefore, these results conclusively prove that the proposed TAR can realistically reflect the geometric stability between adjacent images in MST generation. On the other hand, we can find many image pairs with high OAR and low TAR values from the scatter plot, as shown in Figure 12. This indicates that there are actually many image pairs with tiepoint distribution biased in a wide overlapping region [16]. In addition, we can see that the NoT-based optimal image pairs had the low OAR range. In contrast, the OAR-based optimal image pairs had the low NoT range, while the TAR-based optimal image pairs showed a balanced result between NoT and OAR. These results conclusively demonstrate our assumption that TAR can simultaneously consider both overlap and tiepoint distribution. These results were also confirmed visually by the distance errors shown in Figure 13, where only distance errors larger than 50 pixels are displayed.

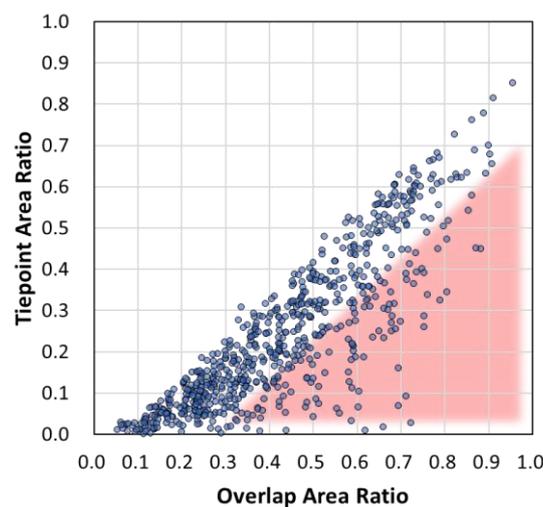


Figure 12. Scatter plot between OAR and TAR values for adjacent images in Dataset-2.

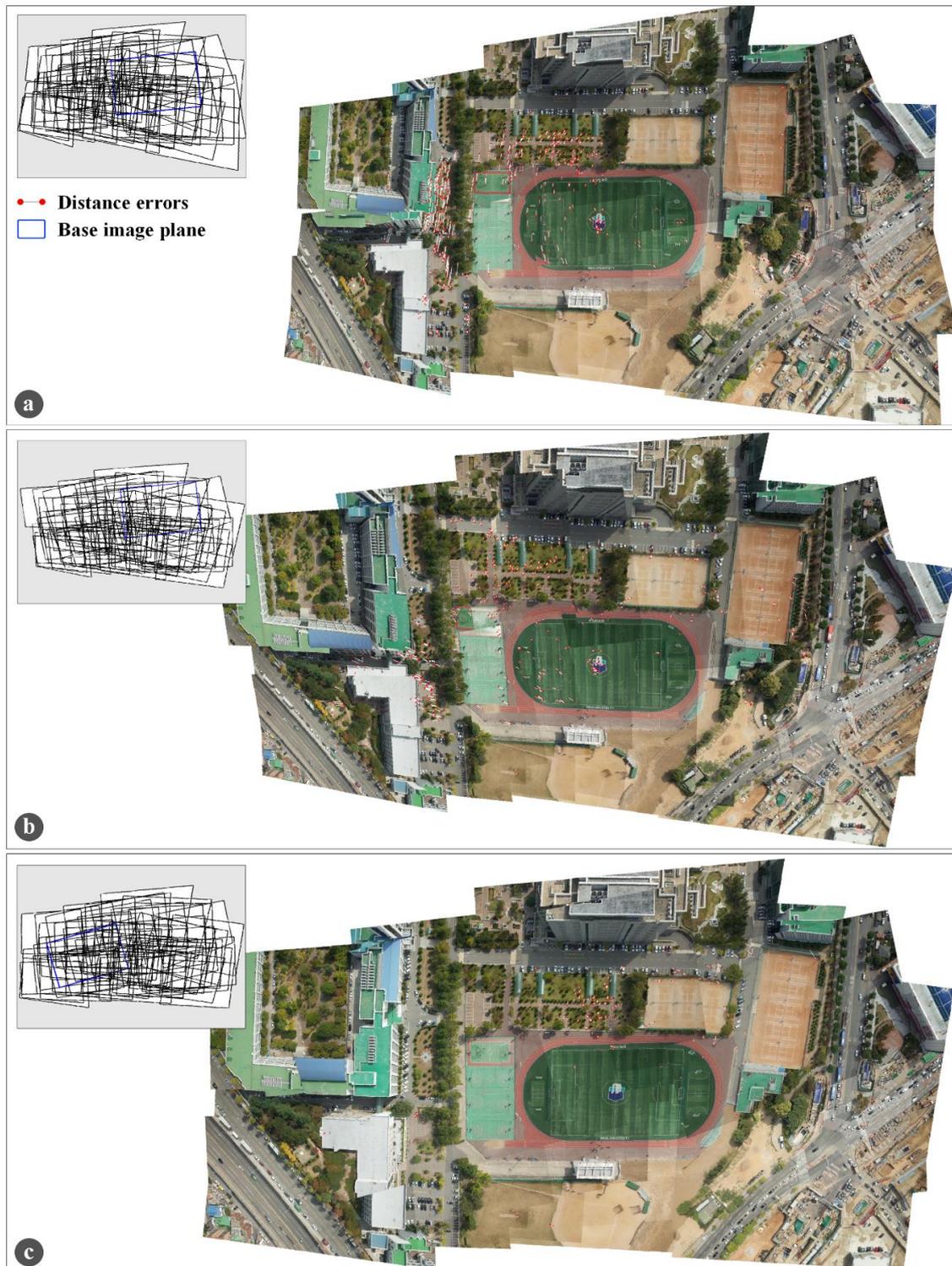


Figure 13. Mosaicking results for a general imaging environment. For pairwise transformation estimation, (a,b) used the homography model and (c) used the hybrid transformation model. For MST generation and mosaic plane selection, (a) applied NoT, (b) applied OAR, and (c) applied TAR and orthogonality.

The proposed method also showed the best performance in terms of mosaic distortion. The mosaicking result from the proposed method produced the smallest amount of deformation, compared with the reference result in Figure 11. This result demonstrates the effectiveness of the

proposed method in mosaic plane selection. The comparative methods generated MSTs with NoT or OAR as a weight and then determined mosaic planes from root images that minimize the depth of the generated MST [21,23,28–30]. This approach may be reasonable, in that the image with the highest geometric stability between adjacent images is set as a mosaic plane [21]. However, this may not take into account the imaging characteristics of small UAVs that are sensitive to environmental changes. Accordingly, the existing methods may determine a relatively tilted image as a mosaic plane [9,13]. This concern was actually realized, as shown in Figure 13a,b.

The evaluation results for the poor imaging environment are summarized in Table 7. In this experiment, the MSTs generated were the same for all methods because overlaps among the images were generally small. This can be seen in that the cases with the same transformation model yielded the same global error. Therefore, this experiment for the poor imaging environment focused on the performance of pairwise transformation modeling.

Table 7. Mosaicking performance evaluations for a poor imaging environment.

Pairwise Modeling	Global Modeling		NoT			OAR			TAR			Reprojection Errors (Pixels)		Distort. (deg.)
	MST Weight	Base Plane	Min.	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.	Pairwise	Global	
Affine	NoT	NoT	12	157	290	0.09	0.37	0.61	0.01	0.24	0.48	27.87	37.57	10.12
Homo.	NoT	NoT	12	157	290	0.14	0.41	0.63	0.01	0.24	0.48	40.24	43.38	25.61
Affine	OAR	OAR	12	157	290	0.09	0.37	0.61	0.01	0.24	0.48	27.87	37.57	10.12
Homo.	OAR	OAR	12	157	290	0.14	0.41	0.63	0.01	0.24	0.48	40.24	43.38	25.61
Hybrid	TAR	Ortho.	12	157	290	0.09	0.37	0.63	0.01	0.24	0.48	22.39	19.56	9.98

The experiment results showed that the affine transformation model can provide better performance than the homography model. This is in contrast to the previous experiment results, which demonstrated our assumption that a simple transformation model would yield better results than a precision transformation model for images with poor geometric stability. Meanwhile, the proposed method produced the best performance again. This means that the proposed method could variably apply optimal transformation models through hybrid transformation modeling, and that the derived TAR value of 0.3 is valid as a criterion for optimal transformation model selection. This criterion is expected to be used in general because it was derived from independent image datasets.

The mosaicking result from the proposed method was better than the comparative methods, both quantitatively and qualitatively. The homography-based results produced large distortions and inconsistencies in the outer images with small overlaps, as seen in the red circles in Figure 14b, and the affine transformation-based results showed some inconsistencies in regions where multiple images were overlaid, as seen in the red circle in Figure 14a.

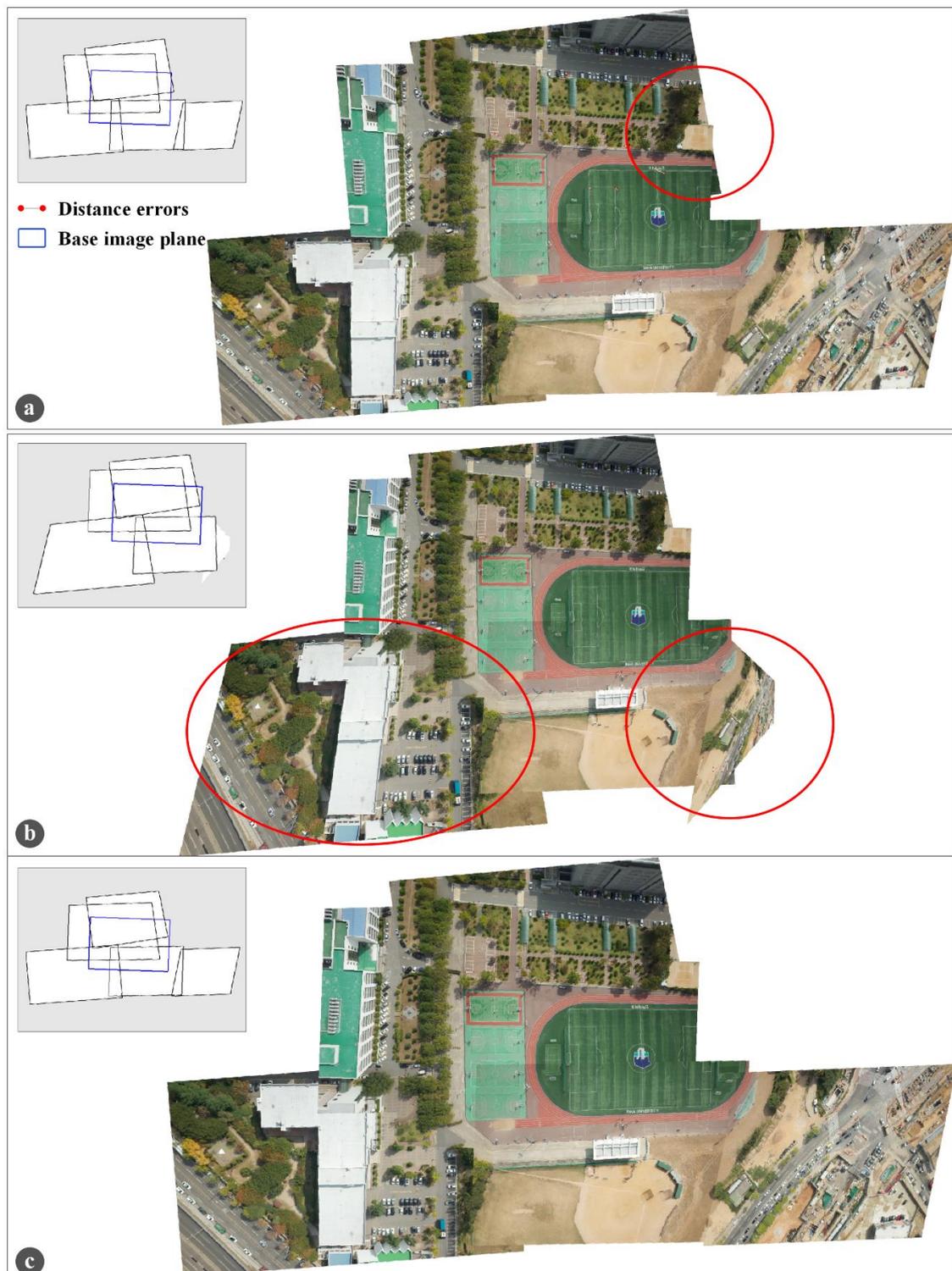


Figure 14. Mosaicking results from a poor imaging environment. For pairwise transformation estimation, (a) used the affine transformation model, (b) used the homography model, and (c) used the hybrid transformation model. For MST generation and base plane selection, (a,b) applied NoT and (c) applied TAR and orthogonality.

4. Conclusions

We developed a robust image mosaicking method that can handle problems caused by the imaging–geometry characteristics of small UAVs. In this paper, the problems were defined as

insufficient overlaps and tilted images owing to unstable flight attitudes and biased tiepoint distributions from low-altitude flights. The proposed method estimated pairwise transformations with optimal transformation models selected by geometric stability analysis between adjacent images. As a geometric stability indicator, TAR was introduced to consider both overlap and tiepoint distribution simultaneously. The valid criterion for the TAR was found to be about 0.3, based on experiments with two independent image datasets. After pairwise transformation modeling between adjacent images, the proposed method estimated global transformations from the MST generated by TAR analysis and the mosaic plane selected by orthogonality analysis. The experiment results showed that the problems raised in this paper could actually occur in image datasets obtained by small UAVs and showed that the proposed method can reliably produce image mosaics for two types of image dataset obtained from general and from extreme imaging environments.

The proposed method does not require any prerequisites in image acquisition, nor any user interventions in image mosaicking. These advantages would even make it possible to mosaic UAV images obtained from a manual flight without the support of a global navigation satellite system. Accordingly, the proposed method could be widely used to quickly and correctly identify situations in sites where the use of existing spatial data and direct access are limited, such as disaster and polar regions. Meanwhile, TAR as proposed in this paper was found to be very effective in geometric stability evaluation between adjacent images. The identification of geometric stability is an important issue in many multiple image-processing techniques, such as structure-from-motion (SfM). Thus, TAR itself also has significant potential for improving many applications in photogrammetry and computer vision.

Author Contributions: Conceptualization, J.-I.K.; Formal analysis, J.-I.K., H.-c.K. and T.K.; Methodology, J.-I.K.; Supervision, H.-c.K. and T.K.; Writing—original draft, J.-I.K.; Writing—review & editing, J.-I.K. and T.K. All authors have read and agreed to the published version of the manuscript.

Funding: The work in this paper was supported by “Cooperative Research Program for Agriculture Science and Technology Development (No.PJ01350003)” of Rural Development Administration, Republic of Korea, and by the Korea Polar Research Institute (KOPRI), Grant PE20080 (Study on remote sensing for quantitative analysis of changes in the Arctic cryosphere).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Turner, D.; Lucieer, A.; Watson, C. An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (SfM) point clouds. *Remote Sens.* **2012**, *4*, 1392–1410. [[CrossRef](#)]
2. Jhan, J.; Rau, J.; Huang, C. Band-to-band registration and ortho-rectification of multilens/multispectral imagery: A case study of MiniMCA-12 acquired by a fixed-wing UAS. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 66–77. [[CrossRef](#)]
3. Kim, T.; Im, Y. Automatic satellite image registration by combination of matching and random sample consensus. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1111–1117.
4. Toutin, T. Geometric processing of remote sensing images: Models, algorithms and methods. *Int. J. Remote Sens.* **2004**, *25*, 1893–1924. [[CrossRef](#)]
5. Gao, F.; Masek, J.; Wolfe, R.E. Automated registration and orthorectification package for Landsat and Landsat-like data processing. *J. Appl. Remote Sens.* **2009**, *3*, 033515.
6. Laliberte, A.S.; Goforth, M.A.; Steele, C.M.; Rango, A. Multispectral remote sensing from unmanned aircraft: Image processing workflows and applications for rangeland environments. *Remote Sens.* **2011**, *3*, 2529–2551. [[CrossRef](#)]
7. Brown, M.; Lowe, D.G. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.* **2007**, *74*, 59–73. [[CrossRef](#)]
8. Du, Q.; Raksuntorn, N.; Orduyilmaz, A.; Bruce, L.M. Automatic registration and mosaicking for airborne multispectral image sequences. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 169–181. [[CrossRef](#)]

9. Xu, Y.; Ou, J.; He, H.; Zhang, X.; Mills, J. Mosaicking of unmanned aerial vehicle imagery in the absence of camera poses. *Remote Sens.* **2016**, *8*, 204. [[CrossRef](#)]
10. Ren, X.; Sun, M.; Zhang, X.; Liu, L. A Simplified method for UAV multispectral images mosaicking. *Remote Sens.* **2017**, *9*, 962. [[CrossRef](#)]
11. Zhou, G.; Ambrosia, V.; Gasiewski, A.J.; Bland, G. Foreword to the special issue on unmanned airborne vehicle (UAV) sensing systems for earth observations. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 687–689. [[CrossRef](#)]
12. Yahyanejad, S.; Rinner, B. A fast and mobile system for registration of low-altitude visual and thermal aerial images using multiple small-scale UAVs. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 189–202. [[CrossRef](#)]
13. Moussa, A.; El-Sheimy, N. A fast approach for stitching of aerial images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 769–774. [[CrossRef](#)]
14. Mehrdad, S.; Safdary, M.; Moallem, P. Toward real time UAVs' image mosaicking. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 941–946. [[CrossRef](#)]
15. Faraji, M.R.; Qi, X.; Jensen, A. Computer vision-based orthorectification and georeferencing of aerial image sets. *J. Appl. Remote Sens.* **2016**, *10*, 036027. [[CrossRef](#)]
16. Yang, Y.; Lin, Z.; Liu, F. Stable imaging and accuracy issues of low-altitude unmanned aerial vehicle photogrammetry systems. *Remote Sens.* **2016**, *8*, 316. [[CrossRef](#)]
17. Xiao, X.; Guo, B.; Li, D.; Li, L.; Yang, N.; Liu, J.; Zhang, P.; Peng, Z. Multi-view stereo matching based on self-adaptive patch and image grouping for multiple unmanned aerial vehicle imagery. *Remote Sens.* **2016**, *8*, 89. [[CrossRef](#)]
18. Kim, J.; Kim, T.; Shin, D.; Kim, S. Fast and robust geometric correction for mosaicking UAV images with narrow overlaps. *Int. J. Remote Sens.* **2017**, *38*, 2557–2576. [[CrossRef](#)]
19. Kim, J.; Kim, T. Development of a robust image mosaicking method for small unmanned aerial vehicle. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *42*, 183. [[CrossRef](#)]
20. Xiang, R.; Min, S.; Cheng, J.; Lei, L.; Hui, Z.; Xiaodong, L. A method of fast mosaic for massive UAV images. In Proceedings of the SPIE 9260, Land Surface Remote Sensing II, Beijing, China, 8 November 2014; p. 92603W.
21. Xia, M.; Yao, J.; Xie, R.; Li, L.; Zhang, W. Globally consistent alignment for planar mosaicking via topology analysis. *Pattern Recognit.* **2017**, *66*, 239–252. [[CrossRef](#)]
22. Alexandre, A.; Ortiz, R.; Vanderghenst, P. Freak: Fast retina keypoint. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, Providence, RI, USA, 16–21 June 2012; pp. 510–517.
23. Marzotto, R.; Fusiello, A.; Murino, V. High resolution video mosaicing with global alignment. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; pp. 692–698.
24. Zaragoza, J.; Chin, T.; Brown, M.S.; Suter, D. As-projective-as-possible image stitching with moving DLT. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *36*, 1285–1298. [[PubMed](#)]
25. Kekec, T.; Yildirim, A.; Unel, M. A new approach to real-time mosaicing of aerial images. *Robot. Auton. Syst.* **2014**, *62*, 1755–1767. [[CrossRef](#)]
26. Shum, H.; Szeliski, R. Construction and refinement of panoramic mosaics with global and local alignment. In Proceedings of the International Conference on Computer Vision, Bombay, India, 7 January 1998; pp. 953–958.
27. Sawhney, H.S.; Kumar, R. True multi-image alignment and its application to mosaicing and lens distortion correction. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 235–243. [[CrossRef](#)]
28. Fang, X.; Luo, B.; Zhao, H.; Tang, J.; Zhai, S. New multi-resolution image stitching with local and global alignment. *IET Comput. Vis.* **2010**, *4*, 231–246. [[CrossRef](#)]
29. Elibol, A.; Gracias, N.; Garcia, R. Fast topology estimation for image mosaicing using adaptive information thresholding. *Robot. Auton. Syst.* **2013**, *61*, 125–136. [[CrossRef](#)]
30. Heiner, B.; Taylor, C.N. Creation of geo-referenced mosaics from mav video and telemetry using constrained optimization bundle adjustment. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, 10–15 October 2009; pp. 5173–5178.

31. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
32. Wang, Z.; Chen, Y.; Zhu, Z.; Zhao, W. An automatic panoramic image mosaic method based on graph model. *Multimed. Tools Appl.* **2015**, *75*, 2725–2740. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).