

Article

Direct Georeferencing for the Images in an Airborne LiDAR System by Automatic Boresight Misalignments Calibration

Haichi Ma ¹, Hongchao Ma ^{1,2,*}, Ke Liu ¹, Wenjun Luo ¹ and Liang Zhang ³

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; mahaichi@whu.edu.cn (H.M.); Liu_ke@whu.edu.cn (K.L.); yunitongzai@whu.edu.cn (W.L.)

² Department of Oceanography, Dalhousie University, Halifax, NS B3H 4R2, Canada

³ Faculty of Resources and Environmental Science, Hubei University, Wuhan 430062, China; zhangliang@hubu.edu.cn

* Correspondence: hchma@whu.edu.cn; Tel.: +86-1380-864-8104

Received: 27 July 2020; Accepted: 2 September 2020; Published: 5 September 2020



Abstract: Airborne Light Detection and Ranging (LiDAR) system and digital camera are usually integrated on a flight platform to obtain multi-source data. However, the photogrammetric system calibration is often independent of the LiDAR system and performed by the aerial triangulation method, which needs a test field with ground control points. In this paper, we present a method for the direct georeferencing of images collected by a digital camera integrated in an airborne LiDAR system by automatic boresight misalignments calibration with the auxiliary of point cloud. The method firstly uses an image matching to generate a tie point set. Space intersection is then performed to obtain the corresponding object coordinate values of the tie points, while the elevation calculated from the space intersection is replaced by the value from the LiDAR data, resulting in a new object point called Virtual Control Point (VCP). Because boresight misalignments exist, a distance between the tie point and the image point of VCP can be found by collinear equations in that image from which the tie point is selected. An iteration process is performed to minimize the distance with boresight corrections in each epoch, and it stops when the distance is smaller than a predefined threshold or the total number of epochs is reached. Two datasets from real projects were used to validate the proposed method and the experimental results show the effectiveness of the method by being evaluated both quantitatively and visually.

Keywords: direct georeferencing; Calibration; digital camera; LiDAR

1. Introduction

Airborne laser scanning, also termed airborne Light Detection and Ranging (LiDAR), is an active remote sensing technique for acquiring 3D geospatial data over the Earth's surface [1,2]. A typical airborne LiDAR system consists of a GPS (Global Positioning System), an IMU (Inertial Measurement Unit), and a laser scanner, with which a point cloud dataset encoding 3D coordinate values under a given geographic coordinate system can be generated [3]. The point cloud can be further processed to extract thematic information and geo-mapping products, such as manmade objects [4], stand-alone plants [5], DEM (Digital Elevation Model)/DTM (Digital Terrain Model) [6], etc. However, there are still many challenges in terms of object detection, extraction, and reconstruction by using the LiDAR dataset alone, because the point cloud provided by a LiDAR system is unstructured, irregularly spaced, and lacks spectral and textural information. Thus, a commercial airborne LiDAR system usually integrates a high-resolution metric digital camera, from which high-resolution aerial images can be collected while collecting point cloud data. The individual characteristics of LiDAR point cloud

and image data are considered complementary [7]. They have been used to enhance the extraction of thematic information by fusing the two datasets for a variety of applications, such as buildings detection and reconstruction [8,9], land cover classification [10,11], road modeling [12,13], and tree species classification [14,15], to name but a few.

In photogrammetric applications, it is necessary to determine the geometric model of the sensing system before the collected images can be used for highly accurate measurement purposes. In traditional aerial photogrammetric mapping, the process begins with the determination of the IOEs (Interior Orientation Elements) and the EOE (Exterior Orientation Elements) of the camera. IOEs are usually provided by the camera manufacturer [16]. This means that IOEs can be viewed as known variables during the photogrammetric processing. EOE can be processed in two steps (relative and absolute orientation), but simultaneous methods (such as bundle adjustments) are now available in the majority of software packages [17]. A photogrammetric test field with highly redundant photo coverage such as 80% forward overlap and 60% side overlap and accurate ground control points (GCPs) are required in the simultaneous methods [18,19]. With the availability of the combination of GPS/IMU, direct georeferencing becomes possible because the EOE can be derived from an integration of relative kinematic GPS positioning and IMU data by Kalman filtering, which is the case in an airborne LiDAR system integrated with a digital camera.

One of the prerequisites for direct georeferencing of images is the rigid connection between the camera and the IMU in order to keep a strict parallel between the image sensing frame and the IMU body frame, which is hard to achieve and may vary even within a given flight day [20]. Moreover, as the origin of the camera frame cannot be coincident with the projection center of the camera, and the GPS antenna will be on the top of the aircraft, the attitude and positional relation between camera and IMU, known as boresight errors/misalignments, must be determined before direct georeferencing of images can be performed, which includes the determination of three rotational angles and three lever arms, as shown in Figure 1. Level arms can be measured by traditional methods, such as direct measurement with ranging tools, close range photogrammetry [21], and accuracy within one centimeter can be achieved [22]. However, the measurements of the boresight misalignments are far more complicated compared to lever arm measurements because no direct methods exist. Conventionally, they are determined indirectly by using a reference block with known ground control points located within the project area or in a special test field, a process termed as system calibration, because it provides the calibration of other parameters such as focal length. Many research works have been conducted regarding direct georeferencing with the conventional method in the past two decades. Heier et al. [23] showed the postprocessing steps of DMC (Digital Metric Camera) image data to generate virtual central perspective images and gave an overview of the entire DMC calibration. Skaloud et al. [24] and Skaloud [25] conducted a study on the method of GPS/IMU integration to provide exterior orientation elements for direct georeferencing of airborne imagery with more reliability and better accuracy. The operational aspects of airborne mapping with GPS/IMU were analyzed and strategies for minimizing the effect of the hardware integration errors on the process for direct georeferencing were proposed. Heipke et al. [26] discussed the direct determination of the EOE via the combination of GPS and IMU as a complete substitute for aerial triangulation. Jacobsen [27] discussed the direct georeferencing based on restoring the geometric relations of images in a chosen object coordinate system, and the possibility of avoiding using control points by direct sensor orientation with the combination of GPS and IMU. Mostafa et al. [28] argued that boresight misalignments calibration is one of the critical steps in direct georeferencing for geomapping purposes. They presented the experimental results of boresight misalignments calibration by using a software and checked the results with ground control points. Honkavaara [29,30] discussed block structures for calibration that significantly affected the cost and efficiency of the system calibration. The experiments indicated that boresight misalignments and the IOEs are the main factors influencing the final results. Jacobsen [31] investigated the direct and integrated sensor orientation based on the combination of relative kinematic GPS and IMU. The investigation showed the advantages of using direct sensor orientation for image

georeferencing without ground control points and independent of block or strip configurations. Filho et al. [32] presented an in-flight calibration method for multi-head camera systems, and the applications of direct georeferencing were evaluated.

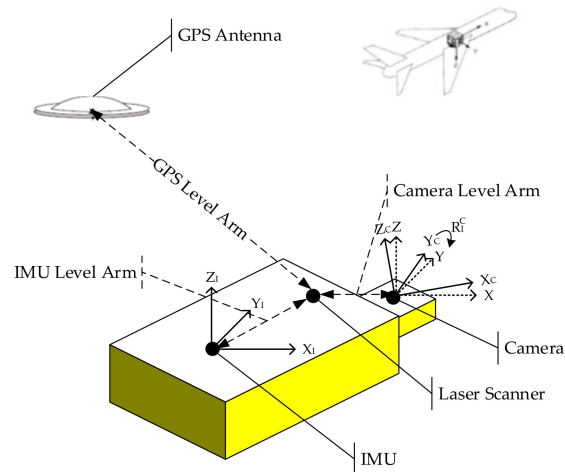


Figure 1. Positional Relationship between Laser Scanner, Camera, GPS, and IMU.

Though intensively adopted in practice, traditional system calibration shows the following drawbacks: Firstly, the environmental conditions such as temperature, humidity, etc. between the test field and mapping areas may dramatically differ. Therefore, the camera geometry during operation may also change relative to the situation in the test field due to changes in environmental conditions [33,34]. Secondly, establishing a new test field for every mapping project and collecting large numbers of ground control points is expensive and sometimes impractical. On the other hand, airborne LiDAR systems deliver direct dense 3D measurements of object surface at a high rate of accuracy [7,16]. Moreover, continued improvements in the performance and accuracies of LiDAR systems in recent years have enabled the use of LiDAR data as a source of control information suitable for photogrammetric applications. Different methods have been tested and implemented for integrating LiDAR and photogrammetric data in performing aerial triangulation or determining the boresight misalignments for direct georeferencing, as will be reviewed in the following.

Delara et al. [35] presented a method to perform the bundle block adjustment using aerial images and laser scanner data. In the method, LiDAR control points were extracted from LiDAR intensity images for determining the exterior orientation elements of a low-cost digital camera. Habib et al. [36,37] utilized linear features derived from LiDAR data as control information for image georeferencing. However, a large number of linear features with good spatial distribution are needed to achieve high accuracy. Kwak et al. [38] proposed using the centroid of the plane roof surface of a building as control information for estimating exterior orientation elements of aerial imagery and registering the aerial imagery relative to the aerial LiDAR data. In the method, the centroid of the plane roof is extracted from aerial imagery by using the Canny Edge Detector and from aerial LiDAR data by using Local Maxima Filtering. Liu et al. [39] presented a method for utilizing LiDAR intensity images to collect high accuracy ground coordinates of GCPs for aerial triangulation process. Yastikli et al. [40] investigated the feasibility of using LiDAR data for in situ calibration of the digital camera. In addition, the determination of attitude and positional relationship between digital camera and IMU was also discussed. Mitishita et al. [41] presented a method of georeferencing photogrammetric images using LiDAR data. The method applied the centroids of rectangular building roofs as control points in the photogrammetric procedure. Ding et al. [42] utilized the vertical vanishing point in an aerial image and the corner points of the roof edge from the point cloud to estimate the pitch and roll of the cameras rotation angles. Based on Ding's study, Wang and Neumann [43] introduced a new feature 3CS (three connected segments) to replace the vanishing point to optimize the method. Each 3CS has three

segments connected into a chain. Wildan et al. [44] utilized control points derived from LiDAR data to perform the aerial triangulation of a large photogrammetric block of analogue aerial photographs. According to the authors, the mapping has achieved the national standard of cartographic accuracy for the 1:50,000 scale mapping. Chen et al. [45] proposed a new method for boresight misalignments calibration of the digital camera integrated in an airborne LiDAR system without ground control points. In the calibration, tie points in overlapping areas are selected manually, and the ground points corresponding to these points are calculated using a multi-baseline space intersection and DEM elevation constraints. Gneeniss [46] and Gneeniss et al. [16] conducted studies on cross-calibrate aerial digital cameras via the use of complementary LiDAR data. The effect of the number and spatial distribution of LiDAR control points to perform aerial triangulation of large photogrammetric blocks was investigated as well.

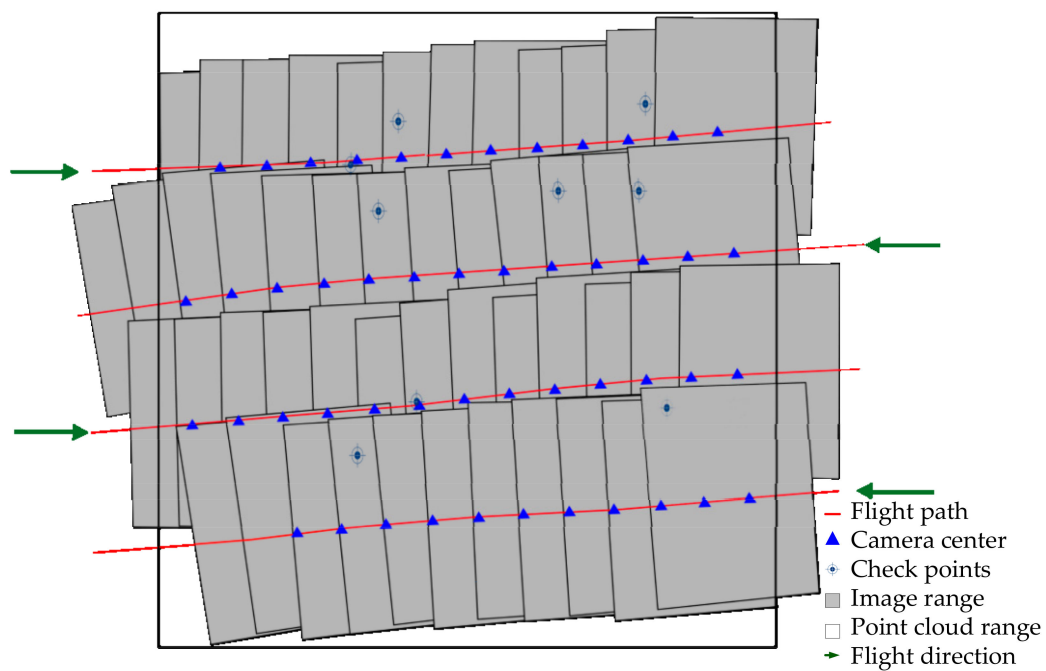
Direct georeferencing of images based on LiDAR point cloud was also provided by commercial software Terrasolid in the TMatch module. In this module, a filter is used firstly to obtain ground points from point cloud, while a large number of tie points of the images are manually selected. The optimal camera misalignment values (new heading, roll, and pitch) are calculated using the tie points from the overlapping images and their corresponding ground points from the LiDAR point cloud. However, image points and LiDAR-derived ground points are matched manually, where artificial errors are inevitable, and matching is impractical when the surveying area is large.

The objective of this study is to introduce a new automatic boresight misalignments calibration method for direct georeferencing of images collected by a digital camera integrated in an airborne LiDAR system. Because the three lever arms can be accurately measured, we focus on the determination of the three rotational angles by using the LiDAR point cloud as auxiliary data. In contrast to the methods presented in previous literature, the in situ camera calibration focuses on using VCPs (Virtual Control Points—defined in following section) and a small sub-block of images selected from the entire block covering the surveying area. The method establishes the error equation by minimizing the distances between the initially selected tie points in the image space and the image points corresponding to VCPs by space resection. The main advantages of the method can be summarized as follows: Firstly, no dedicated calibration test fields, or even ground control points, are needed. Secondly, the whole procedure is fully automatic, from the extraction of tie points to the calibration of the boresight misalignments. This is of particular importance when the airborne LiDAR system is employed to collect data for rapid response to natural disasters, such as earthquake relief efforts. Finally, the accuracy of the georeferenced images is high enough for many geospatial applications, as shown in the experimental results.

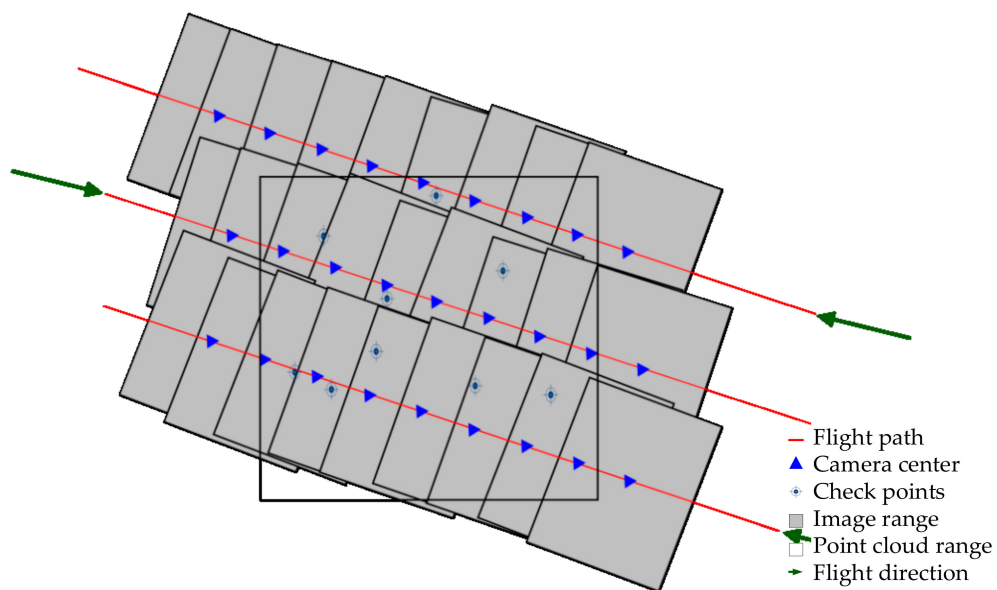
2. Materials and Methods

2.1. Sample Materials

Two datasets were used to test the effectiveness of the proposed method. The first dataset was acquired in a suburb of Xi'an, Shaanxi province, China, and consists of LiDAR point cloud and aerial images. Images were collected by Leica RCD105 and point cloud was acquired by Leica ALS60. Specifications of the equipment and the data are listed in Table 1. Figure 2a shows the flight direction of the strips. The selected sub-set this dataset covers extends approximately 3.26 km in the east–west and 3.34 km in the north–south direction. The second dataset came from Ningbo, Zhejiang Province, China, and also consists of point cloud and optical images. Although both of the LiDAR systems adopted to collect the two datasets were manufactured by Leica Geosystem, specifications of the equipment and data are different (see Table 2 and Figure 2b for more details of the second dataset).



(a)



(b)

Figure 2. Block layout configuration: (a) Dataset 1; and (b) Dataset 2.

Outliers in the point cloud were filtered before it was input to the calibration process. This can be achieved either by a commercial software or open source libraries such as PCL (Point Cloud Library) [47].

Table 1. Technical parameters of Dataset 1.

| Flight Information | Target Area | Max Flight Height | Min Flight Height | Number of Flights | Number of Images |
|--------------------|----------------|---------------------|-------------------|--------------------|------------------|
| | Xi'an, China | | 1450 m | 1387 m | 4 |
| LiDAR Points | Sensor | Point Cloud Density | FOV | Size of Area | Average Overlap |
| | Leica ALS60 | 1.9/m ² | 45° | 11 km ² | 48% |
| Aerial Images | Type of Camera | Pixel Size | Focal Length | Forward Overlap | Side Overlap |
| | RCD105 | 0.0068 mm | 60 mm | 70% | 45% |

Table 2. Technical parameters of Dataset 2.

| Flight Information | Target Area | Max Flight Height | Min Flight Height | Number of Flights | Number of Images |
|--------------------|----------------|---------------------|-------------------|--------------------|------------------|
| | Ningbo, China | | 2397 m | 2304 m | 3 |
| LiDAR Points | Sensor | Point Cloud Density | FOV | Size of Area | Average Overlap |
| | Leica ALS70 | 0.60/m ² | 45° | 15 km ² | 20% |
| Aerial Images | Type of Camera | Pixel Size | Focal Length | Forward Overlap | Side Overlap |
| | RCD30 | 0.006 mm | 53 mm | 60% | 30% |

2.2. Methods

2.2.1. Overview of the Methods

We first define the concept of virtual control points. For an airborne LiDAR system integrated with a digital camera, assuming that boresight misalignments have been corrected and lever arms were accurately measured, select a sub-block of images where two adjacent images meet the requirement of at least 60% forward overlap. Then, for any pair of tie points in the two images, their object coordinates can be derived from the collinearity equations with the exterior orientation elements provided by GPS/IMU combined navigation system, or synonymously termed as Positioning and Orientation System (POS) in the airborne LiDAR community, which are accurate enough because boresight misalignments were corrected and lever arms were accurately measured. Thus, a given pair of tie points

a_1 and a_2 (see Figure 3), determine a corresponding object point, denoted by P_{image} with object coordinates (X, Y, Z_{image}) under a given object coordinate system. If the adjacent images cover a flat area, then the elevation value of the area shall be a constant, which can be derived from LiDAR point cloud and be denoted by Z_{laser} . Currently, vertical accuracy better than several centimeters can be achieved in flat areas by a commercial airborne LiDAR system [48]; thus, in most topographic mapping applications, the elevation value from the point cloud can be treated as the true value. If a_1 and a_2 were accurately matched, the object point P_{image} should lie in the flat plane. However, due to the existence of systematic errors, such as boresight misalignments and other random errors, P_{image} will be off the plane; that is, its elevation value Z_{image} can be significantly greater or less than the elevation value derived from point cloud Z_{laser} . Replace Z_{image} by Z_{laser} to create an object point with object coordinates (X, Y, Z_{laser}) , which is defined as a Virtual Control Point (VCP).

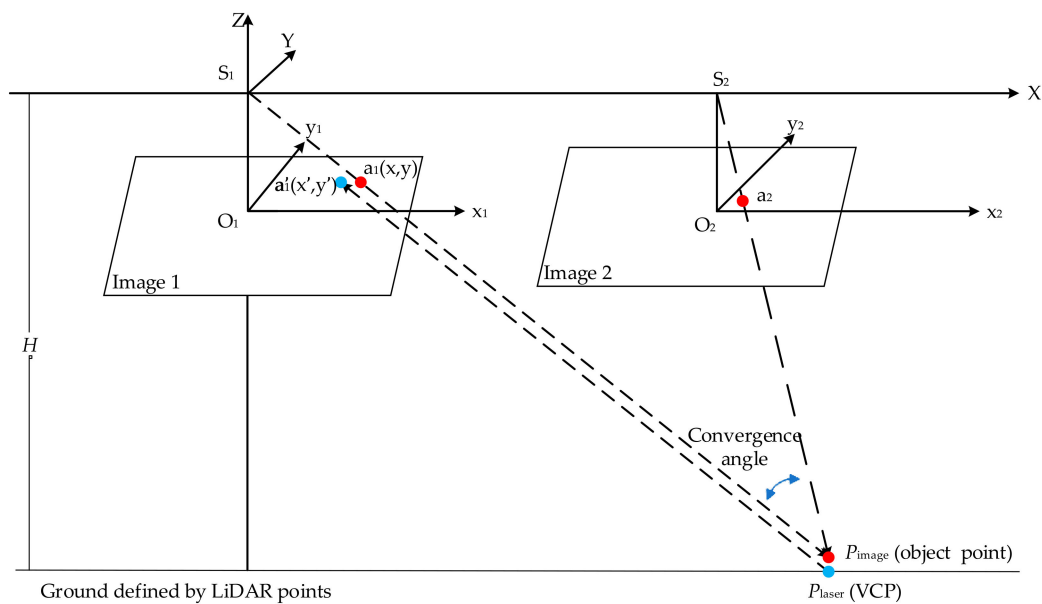


Figure 3. The distance between a_1 and a'_1 : S_1 and S_2 are the spatial positions of the projection center of the camera at the moment of acquiring Images 1 and 2. H is the altitude of the aircraft.

Our method begins with image matching to generate a set of tie points. For each pair of tie points, a VCP can be derived as described above. As shown in Figure 3, denote the image coordinates of tie point a_1 by x and y and denote its VCP by P_{laser} . Reproject P_{laser} onto Image 1 by using collinearity equations, and then an image point a'_1 can be obtained. If $Z_{\text{laser}} = Z_{\text{image}}$, then a_1 and a'_1 are identical. However, due to the existence of boresight misalignments and other random errors, the distance between a_1 and a'_1 cannot be neglected. The total distance between all of the tie points and their corresponding points that are derived from VCPs is calculated. Minimizing the total distance by iteratively correcting the boresight misalignments is the core idea of the proposed method.

The general workflow of the method is illustrated in Figure 4. The main steps include:

1. Select sub-block images collected over a relatively flat area from image set and extract tie points in the overlapping areas of the sub-block images using a Speed-Up Robust Features (SURF) algorithm [49].
2. For each pair of tie points, the object point can be derived by space intersection.
3. Replace the elevation values of object points by those derived from LiDAR point cloud. These new points are called VCPs.
4. An automatic VCP selection procedure is designed to perform various assessments to guarantee high quality VCPs can be selected.
5. Adjustment equation is established to perform boresight misalignments compensation based on the combination of the VCP set and collinearity equations.
6. Repeat Steps 2–5 until the total distance between all tie points and their corresponding points that are derived from VCPs remains stable in the iteration or the maximum iteration has been reached.

In the following subsection, detailed explanations are stated, and key equations and formulas are provided.

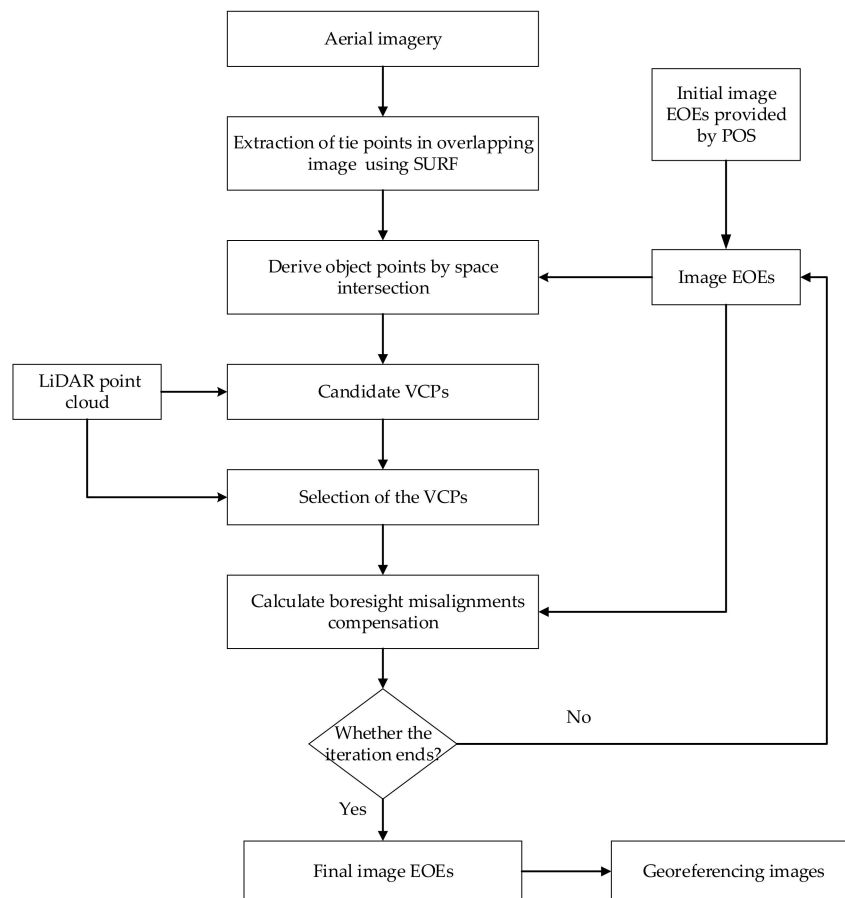


Figure 4. Flowchart of the proposed method.

2.2.2. Detection and Matching of Tie Points in Overlapping Images Using SURF Algorithm

Detection and matching tie points in overlapping image areas is the first step in the method. The removal of boresight misalignments is largely affected by the accuracy of tie points. Thanks to the progress of image matching techniques [49–54], automatic tie points extraction with high accuracy becomes operational.

Although there are many image matching algorithms available, Speed-Up Robust Feature (SURF) [49] is adopted in this study. It is a robust local feature point descriptor, which consists of three major stages: (1) interest point detection; (2) orientation assignment; and (3) interest point description. In the first stage, potential interest points are identified by scanning the image over location and scale. This is effectively achieved by the Hessian matrix using speckle detection. In the second stage, the dominant orientation for each interest point is identified based on its local image patch. The third stage builds a local image descriptor for each interest point, based upon the image gradients in its local neighborhood. This is followed by establishing correspondences with interest points of different images to obtain connection points.

The performance of the SURF algorithm is similar to Scale Invariant Feature Transform (SIFT) [51] in many respects, including robustness to lighting, blur, and perspective distortion. However, it improves computational efficiency by using integral images, Haar wavelet transforms, and approximate Hessian matrix operations. In direct georeferencing, the faster the calculation speed is, the higher the value is in practical survey tasks, without sacrificing an accurate performance. Due to its high efficiency, accuracy, and robustness, we use the SURF algorithm to extract image connection points. In addition, in the process of feature point registration, as the image has been positioned using the initial elements provided by the POS system, only the feature points in the center S -scale window where the feature

points are located are paired. This method can effectively improve computing efficiency, as shown in the experiment section of the study.

For each pair of tie points, the object point is derived by space intersection where exterior orientation elements are provided directly by the POS system without the correction of boresight misalignments.

The fundamental of space intersection is that conjugate light rays meet in the object space. The angle between two conjugate rays is defined as convergence angle in [55] (see Figure 3), and the authors showed that the positional accuracy decreases dramatically when the convergence angle is less than 14.6 degrees. Therefore, if the convergence angle is less than this threshold, the object point obtained by space intersection is labeled as a non-candidate; otherwise, it is labeled as a candidate from which the VCPs will be selected. The convergence angle can be calculated according to Equation (1).

$$\begin{aligned} \cos \varphi &= \frac{\vec{S_1a_1} \cdot \vec{S_2a_2}}{|\vec{S_1a_1}| \cdot |\vec{S_2a_2}|} \\ &= \frac{(X_{s_1} - X_{a_1}) \cdot (X_{s_2} - X_{a_2}) + (Y_{s_1} - Y_{a_1}) \cdot (Y_{s_2} - Y_{a_2}) + (Z_{s_1} - Z_{a_1}) \cdot (Z_{s_2} - Z_{a_2})}{\sqrt{(X_{s_1} - X_{a_1})^2 + (Y_{s_1} - Y_{a_1})^2 + (Z_{s_1} - Z_{a_1})^2} \cdot \sqrt{(X_{s_2} - X_{a_2})^2 + (Y_{s_2} - Y_{a_2})^2 + (Z_{s_2} - Z_{a_2})^2}} \end{aligned} \quad (1)$$

where S_1 and S_2 are the center positions of the camera at the moment of acquiring Images 1 and 2, respectively; and a_1 and a_2 are a pair of tie points.

2.2.3. Selection of the VCPs

A large number of tie points can be detected from two overlapping images after the tie points detection step. Theoretically, the same amount of object points corresponding to tie points can be derived by space intersection.

The first step of VCP selection is the replacement of the elevation values of object points by those derived from LiDAR point cloud. This is a relatively simple step, which mainly involves searching the point cloud via the x and y coordinates of the object points and then replacing their elevation values by interpolation; a new point set called candidate VCPs is obtained, from which the VCPs will be selected. In detail, the step proceeds as follows: the point cloud is projected onto the x–y plane, and then a 2D triangulated irregular network (TIN) is constructed based on the projected points. For a given object point (X, Y, Z) , it is determined which triangle the candidate falls into. If it happens to be positioned at one vertex of a triangle, then the Z value is replaced by that vertex. Otherwise, it is estimated by inverse distance weighted interpolation with elevation values of the three vertices as known points to the interpolation (see Figure 5).

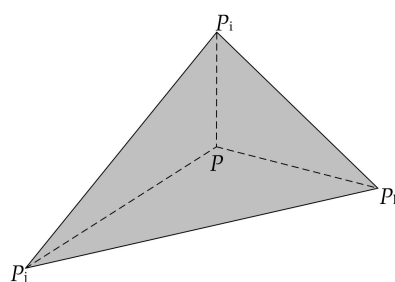


Figure 5. A VCP point P inside a triangle formed by P_i , P_j , and P_k .

Importantly, as defined in Section 2.2.1, a VCP should be located in a flat patch of a building or a terrain surface so that the elevation values of the object points in it are insensitive to planar positions; that is, given two object points $P_1 (X_1, Y_1, Z_1)$ and $P_2 (X_2, Y_2, Z_2)$ in a flat patch, then elevation values $Z_1 \approx Z_2$ regardless of where P_1 and P_2 are located in the patch. This is a very important constraint of the proposed method because it makes the method insensitive to the planar errors of the object points, which is mainly caused by inaccurate image matching. Thus, several criteria for VCP selection have

been designed (see below) in order to guarantee that all the VCPs are located in flat or approximately flat patches and that all of them are reliable.

Planarity: Flat patches, roads, playgrounds, flat building roofs, etc. are the main surfaces where VCPs can be located. Planarity is measured by finding the best-fit plane for the LiDAR points in a window with a predefined size centered at a VCP. At least four points are required for a plane fitting (Equation (2)), thus the window size and the density of the point cloud are two key parameters for the plane fitting. In [56], a more specific definition of density for building recognition is proposed and it is pointed out that an average density of 1 point/m² can detect a building roof with size 2.8 m × 2.8 m. Thus, considering that in most practical cases the average density is higher than 1 point/m², a 3 m × 3 m size window is adopted in the study.

$$A \times x + B \times y + C \times z + D = 0 \quad (2)$$

A threshold of 0.2 m was set to check if a candidate VCP is located in the fitted plane: if the distance from it to the fitted plane less than 0.2 m, then it remains in the VCP set; otherwise, it is labeled as non-VCP.

Slope tolerance: An absolute flat patch is almost nonexistent in real situations, thus it is valuable to study the elevation variation with the change of the slope of a planar patch along the steepest ascend or descend direction (see Figure 6). A simple calculation shows that, when the distance between P_1 and P_2 is 1 m, then the elevation difference is 14 cm when the slope is 8 degrees. Bearing in mind that most commercial LiDAR systems achieve vertical accuracy better than 14 cm, planar patches with slopes less than or equal to 8 degrees were classified as flat patches.

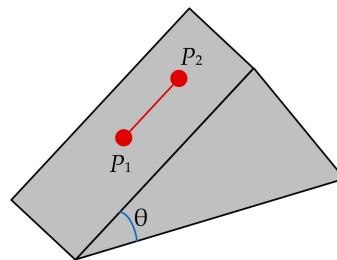


Figure 6. Elevation difference on slope.

Reliability checking: Theoretically, two stereo images reconstruct the 3D terrain of the area covered by the overlapping part of the images. Presently, however, overlapping more than 80% is common in practice thanks to the development of digital cameras, which allow more than two images to overlap the same area. Multiple stereo images not only improve the accuracy of the object coordinates of tie points, but also provide a method for reliability checking of the candidate VCPs. This is because the errors along the x -axis of an object point intersected by two rays cannot be readily detected, which can be overcome by the intersection of more than two rays [57]. This check is not required because VCPs are selected from flat patches, which means the planar errors are insensitive to the elevation values and that the process of boresight misalignments correction is an iterative one, leaving the planar errors to be corrected in the iterations.

2.2.4. Boresight Misalignments Calibration

As stated in the introduction, there is much literature available regarding boresight misalignments calibration. A new calibration method was proposed by Baumker and Heimes in 2002 [58], in which the boresight misalignments were contained in a misalignments matrix, which, with other transformation matrices, forms a complete transformation matrix from terrain to image system. An adjustment equation was then formed for each photo, followed by the establishment of a total normal equation including all measurements. Straying from Baumker's method, our model is constrained by the

minimization of the total distance between tie points and image points calculated from VCPs by space resection, thus the transformation matrix containing the boresight misalignments is included in the colinear equations where boresight misalignments are treated as extra rotations around three axes alongside three exterior orientation elements, $\varphi, \omega,$ and k . Figure 7 shows the coordinate transformation for direct georeferencing.

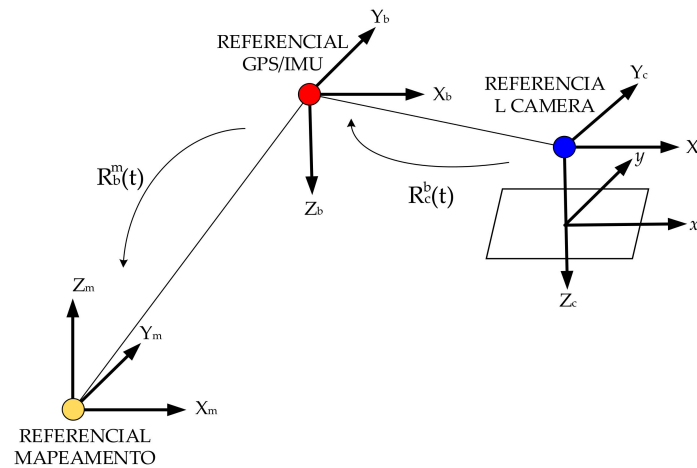


Figure 7. Coordinate transformation for direct georeferencing.

The proposed method is based on the traditional collinear equations, which relate the relationship among image pixels coordinates, ground points coordinates, and photographic center as shown in the following equations:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda R_b^m(\varphi, \omega, \kappa) \begin{bmatrix} x \\ y \\ -f \end{bmatrix} + \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} = \lambda \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ -f \end{bmatrix} + \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix}, \quad (3)$$

where

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{bmatrix} \begin{bmatrix} \cos \kappa & -\sin \kappa & 0 \\ \sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

The colinear equations taking the boresight misalignments transformation matrix into account can be expressed as Equation (5):

$$\begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} = \lambda R_b^m(\varphi, \omega, \kappa) R_c^b(\varphi', \omega', \kappa') \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = \lambda \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} a_1' & a_2' & a_3' \\ b_1' & b_2' & b_3' \\ c_1' & c_2' & c_3' \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix}, \quad (5)$$

where

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{bmatrix} \begin{bmatrix} \cos \kappa & -\sin \kappa & 0 \\ \sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (6)$$

and

$$\begin{bmatrix} a_1' & a_2' & a_3' \\ b_1' & b_2' & b_3' \\ c_1' & c_2' & c_3' \end{bmatrix} = \begin{bmatrix} \cos \varphi' & 0 & -\sin \varphi' \\ 0 & 1 & 0 \\ \sin \varphi' & 0 & \cos \varphi' \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega' & -\sin \omega' \\ 0 & \sin \omega' & \cos \omega' \end{bmatrix} \begin{bmatrix} \cos \kappa' & -\sin \kappa' & 0 \\ \sin \kappa' & \cos \kappa' & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (7)$$

x and y denote the image coordinates of a tie point in the image coordinate system; f is the focal length of the camera; X, Y, Z are the coordinates of the ground point corresponding to the image pixel in the object coordinate system; X_S, Y_S, Z_S are the coordinates of the projection center of the camera; λ is a scale factor; and a_i, b_i, c_i are the elements of the rotation matrix $R_b^m(\varphi, \omega, \kappa)$. Boresight misalignments are denoted by $\varphi', \omega', \kappa'$, and its rotation matrix is denoted by $R_c^b(\varphi', \omega', \kappa')$.

In general, the values of $\varphi', \omega', \kappa'$ are very small. Thus, the matrix of boresight misalignments can be approximated as follows:

$$R_c^b(\varphi', \omega', \kappa') = \begin{bmatrix} 1 & 0 & -\varphi' \\ 0 & 1 & 0 \\ \varphi' & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\omega' \\ 0 & \omega' & 1 \end{bmatrix} \begin{bmatrix} 1 & -\kappa' & 0 \\ \kappa' & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\kappa' & -\varphi' \\ \kappa' & 1 & -\omega' \\ \varphi' & \omega' & 1 \end{bmatrix} \quad (8)$$

Equation (5) can be compactly rewritten as:

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} a_1' & a_2' & a_3' \\ b_1' & b_2' & b_3' \\ c_1' & c_2' & c_3' \end{bmatrix}^{-1} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}^{-1} \begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix}, \quad (9)$$

or

$$\begin{cases} x - x_0 = -f \frac{m_{11}(X - X_S) + m_{12}(Y - Y_S) + m_{13}(Z - Z_S)}{m_{31}(X - X_S) + m_{32}(Y - Y_S) + m_{33}(Z - Z_S)} \\ y - y_0 = -f \frac{m_{21}(X - X_S) + m_{22}(Y - Y_S) + m_{23}(Z - Z_S)}{m_{31}(X - X_S) + m_{32}(Y - Y_S) + m_{33}(Z - Z_S)} \end{cases}, \quad (10)$$

m_{ij} are the elements in the matrix generated by multiplying the inverse of R_b^m and R_c^b .

Take the x, y coordinates of an image tie point as the observations and boresight misalignments $\varphi', \omega', \kappa'$ as the unknowns. Denote the initial values of x and y by (x) and (y) , respectively, which are calculated from Equation (5), and then the linearization of Equation (10) can be expressed as:

$$\begin{cases} x = (x) + \frac{\partial F_x}{\partial \varphi'} \Delta \varphi' + \frac{\partial F_x}{\partial \omega'} \Delta \omega' + \frac{\partial F_x}{\partial \kappa'} \Delta \kappa' \\ y = (y) + \frac{\partial F_y}{\partial \varphi'} \Delta \varphi' + \frac{\partial F_y}{\partial \omega'} \Delta \omega' + \frac{\partial F_y}{\partial \kappa'} \Delta \kappa' \end{cases} \quad (11)$$

Take the V as the correction value, then the error equation can be established as:

$$V = AB - L, \quad (12)$$

where

$$V = [v_x, v_y]^T, \quad B = [\Delta \varphi', \Delta \omega', \Delta \kappa']^T, \quad A = \begin{bmatrix} \frac{\partial F_x}{\partial \varphi'} & \frac{\partial F_x}{\partial \omega'} & \frac{\partial F_x}{\partial \kappa'} \\ \frac{\partial F_y}{\partial \varphi'} & \frac{\partial F_y}{\partial \omega'} & \frac{\partial F_y}{\partial \kappa'} \end{bmatrix}$$

$$L = [l_x, l_y]^T = [x - (x), y - (y)]^T,$$

A normal equation can be obtained by the indirect adjustment principle:

$$A^T W A B = A^T W L, \quad (13)$$

where W is the weight matrix of the observations, which can be assigned a unit matrix, because all observations are assumed to have the same accuracy. Therefore, the corrections of the boresight misalignments can be calculated as follows:

$$B = (A^T A)^{-1} A^T L \quad (14)$$

It is worth pointing out that, while the three rotation angles φ, ω, κ can be derived from data of the POS system, the angular elements it provides are heading and pitch, which do not comply with rotation angles used in photogrammetry. Several methods were developed to transform these two sets

of angles. We adopted the method proposed by Zhao et al. [59], in which the compensation matrix is not required.

The next step is to calculate the image coordinates (x', y') corresponding to the virtual control points using Equation (5) with the corrected boresight misalignments and other exterior orientation elements of that image.

Assume there are n tie points, for a given point (x_i, y_i) , and calculate the distance between it and its corresponding (x'_i, y'_i) . Then, average error (Ex, Ey) and RMSE (Root Mean Square Error) (Rx, Ry) can be calculated by the following formula:

$$\left\{ \begin{array}{l} dx_i = x_i - x'_i \\ dy_i = y_i - y'_i \\ Ex = \frac{1}{n} \sum_{i=1}^n |dx_i| \\ Ey = \frac{1}{n} \sum_{i=1}^n |dy_i| \\ Rx = \sqrt{\frac{1}{n} \sum_{i=1}^n (dx_i)^2} \\ Ry = \sqrt{\frac{1}{n} \sum_{i=1}^n (dy_i)^2} \end{array} \right. \quad (15)$$

These errors plus the maximum iterations determine when the iteration will be stopped: if all four errors remain stable in the iteration, or maximum iterations have been reached, then the iteration stops and the boresight misalignments are obtained.

3. Results

3.1. Results of Tie Points Detection and Matching

Tie points detection and matching by SURF is described in detail in Section 2.2. SIFT (Scale Invariant Feature Transform) algorithm was adopted for the sake of comparison. Figures 8 and 9 show the results by SIFT and SURF detection and matching, respectively. Table 3 lists the performance of the two algorithms, including the number of tie points detected and matched, the time cost, and the accuracy.

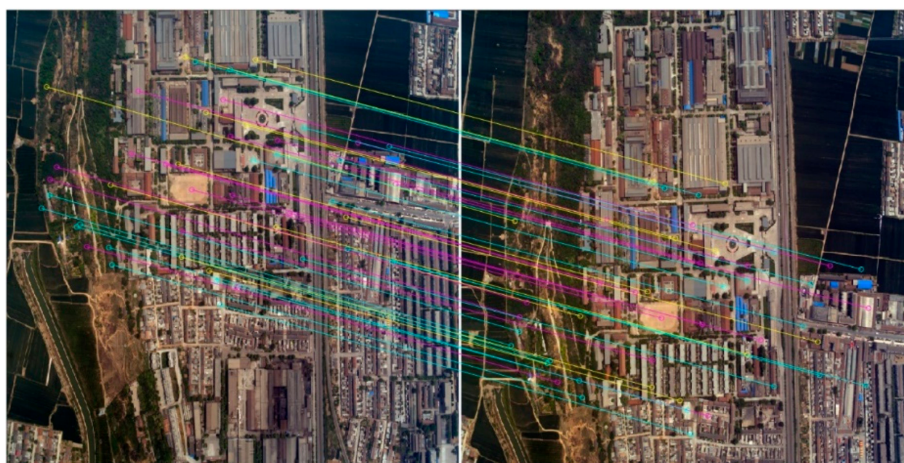


Figure 8. Partial view of the tie points obtained by SIFT (showing only 50 tie points).



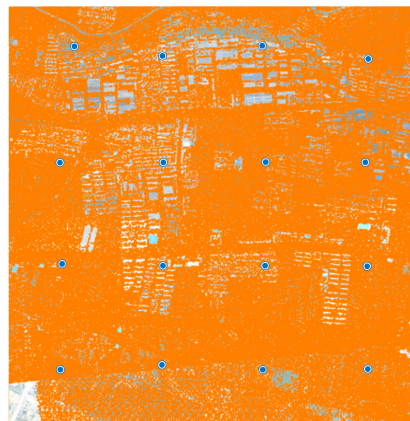
Figure 9. Partial view of the tie points obtained by SURF (showing only 50 tie points).

Table 3. Precision parameter of tie points extraction experiment.

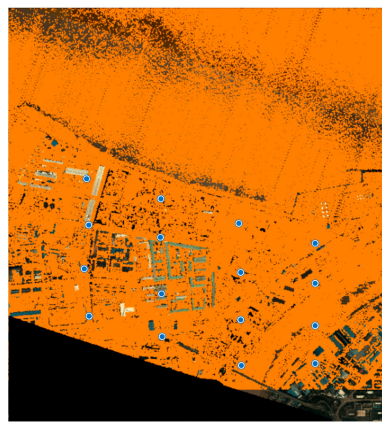
| Algorithm | Tie Points Offset | Time Cost | Number of Tie Points | Correct Tie Points | Accuracy |
|-----------|-------------------|-----------|----------------------|--------------------|----------|
| SURF | 1–3 pixels | 17 s | 1036 | 832 | 80.3% |
| SIFT | 1–3 pixels | 329 s | 2259 | 1830 | 81.0% |

As shown in Table 3, SURF and SIFT have roughly the same tie points offset and matching accuracy. Both can detect and match enough tie points for direct georeferencing purposes. In the experiment, the SURF algorithm detected and matched 1036 tie points, among which 832 are correct, while the SIFT algorithm extracted 2259 tie points in total, and 1830 are correct. However, the time cost of the SIFT algorithm is significantly higher than the SURF algorithm, which coincides with previous comparison studies such as those carried out by Mistry and Banerjee [60]. Thus, we use the SURF algorithm for tie points detection and matching.

Space intersection was performed to the tie point dataset, from which object points corresponding to tie points were obtained, which in turn formed the candidates of virtual control points. Because redundant tie points were detected and matched thanks to the SURF algorithm, so does the object points. A VCP selection procedure described in Section 2.2.3. was applied to select virtual control points. Figure 10 shows the locations of VCPs of the two datasets. Since the upper half area of Dataset 2 covers a beach area, tie points extracted were error prone, and preference was given to the tie points distributed in the lower part of the dataset when space intersection was performed.



(a)



(b)

Figure 10. Virtual control points (16 selected VCPs): (a) Dataset 1; and (b) Dataset 2.

3.2. Results of Direct Georeferencing by Boresight Misalignments Calibration

As described in Section 2.2.4, boresight misalignments were calibrated with VCPs in an iterative manner, followed by direct georeferencing for the two image datasets. The effectiveness of boresight misalignments calibration can be evaluated by assessing the accuracy of the georeferencing images before and after the misalignments were corrected. Other visual inspections of the effectiveness include the check of the mosaiced two adjacent images to see if continuous linear features such as roads are seamlessly mosaiced, and the registration of the point cloud and the georeferenced images. Since accuracies were achieved for the two datasets, quantitative evaluations were given to the first dataset only to avoid a lengthy repetitive analysis. Figures for visual inspection are provided for the second dataset.

The misalignments of the first datasets and their relationship with the number of images used for calibration are tabulated in Table 4. As shown in the table, when the number of images is small and the images are distributed in less strips, the calibration results are unstable; when four strips are used, and the number of images greater than 16, calibration results tend to be stable.

Table 4. Misalignments (in degrees) calibrated by different numbers of images.

| Number of Strips | Number of Images | φ' | ω' | κ' |
|------------------|------------------|------------|-----------|-----------|
| 2 | 2 | 0.3973 | 0.6913 | 0.3122 |
| | 4 | −0.3022 | 0.5170 | 0.1613 |
| | 6 | −0.3014 | 0.5640 | 0.3399 |
| | 8 | −0.3719 | 0.5918 | 0.2766 |
| | 10 | −0.3519 | 0.6018 | 0.2966 |
| 3 | 3 | −0.3767 | 0.6529 | 0.3835 |
| | 6 | −0.3255 | 0.5621 | 0.3029 |
| | 9 | −0.3240 | 0.5587 | 0.2987 |
| | 12 | −0.3320 | 0.5301 | 0.2758 |
| | 15 | −0.3120 | 0.5605 | 0.2951 |
| 4 | 4 | −0.3552 | 0.4682 | 0.3967 |
| | 8 | −0.3417 | 0.5592 | 0.2964 |
| | 16 | −0.3223 | 0.5619 | 0.2957 |
| | 20 | −0.3221 | 0.5615 | 0.2962 |
| | 24 | −0.3222 | 0.5616 | 0.2958 |

The mosaic images of the two datasets after boresight misalignments calibration are shown in Figure 11.

Eighteen check points (their distribution is shown in Figure 7) collected by RTK (Real-Time Kinematic) technique were used to evaluate the planar accuracy of the georeferenced images of the first image dataset. The results are tabulated in Table 5. Six methods published in literature were selected to compare the performance with ours, which was measured by Root Mean Squared Error (RMSE) of the georeferenced images (see Table 6)

Table 5. Planar accuracy of the georeferenced images.

| Residual Errors (m) | Georeferenced Images Before Boresight Misalignments Calibration | | | Georeferenced Images After Boresight Misalignment Calibration | | |
|---------------------|---|--------|--------|---|---------|--------|
| | dX | dY | dXY | dX | dY | dXY |
| Average value | 1.0241 | 2.9382 | 3.1115 | 0.2398 | −0.3511 | 0.5065 |
| Max value | 1.7276 | 5.4239 | 5.6924 | 0.8113 | 1.1249 | 1.3869 |
| RMSE | 0.8943 | 1.6423 | 1.8700 | 0.3915 | 0.5135 | 0.6459 |

Table 6. Comparison of the performance of our method and six others by RMSE along x and y axes, which are denoted RMSE_x and RMSE_y, respectively. $RMSE_{xy} = \sqrt{RMSE_x^2 + RMSE_y^2}$.

| Method from Literature Numbered | RMSE _x (m) | RMSE _y (m) | RMSE _{xy} (m) |
|---------------------------------|-----------------------|-----------------------|------------------------|
| [16] | 0.11 | 0.05 | 0.12 |
| [35] | 0.35 | 0.45 | 0.57 |
| [36] | 0.25 | 0.16 | 0.30 |
| [39] | 0.45 | 1.18 | 1.26 |
| [40] | 0.33 | 0.33 | 0.47 |
| [45] | 0.35 | 0.77 | 0.85 |
| Ours | 0.39 | 0.51 | 0.64 |

As shown in Table 5, the RMSE of the first dataset is approximately 0.39, 0.51, and 0.64 m in X offset, Y offset, and XY offset, respectively, decreasing 2–3 times compared with those before boresight misalignments calibration.

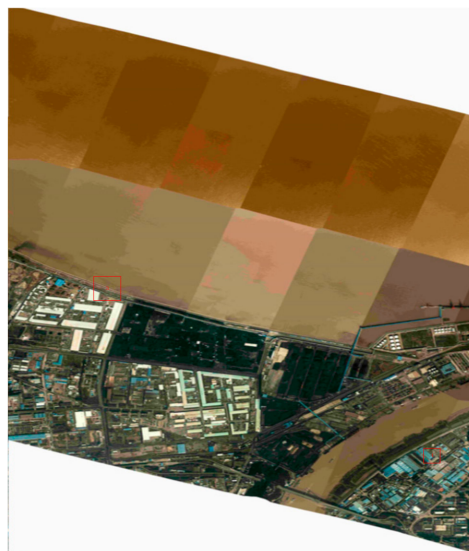
Tables 5 and 6 show that some methods achieve higher accuracy than ours. This is mainly due to the fact that tie points or LiDAR control points are extracted manually by the methods in [35,39,40,45]. While the method in [16] perform two steps before extracting the LiDAR control

points: the determination of the initial coordinates of the photogrammetric point cloud and the registration of the photogrammetric point cloud to the reference LiDAR surface using least squares surface matching method.

Figure 12 provides the visual check for the plane accuracy of the geo-referenced image before and after boresight misalignments calibration (red rectangle in Figure 11), while Figure 13 shows the result of LiDAR point cloud overlapped with georeferenced images. Together with the quantitative accuracy evaluation, it is obvious that the proposed method is effective, and has a great value in case of specific applications such as rapid response where no control points could be collected and where relatively lower accuracy is required.



(a)



(b)

Figure 11. A panoramic view of the DOM generated with boresight misalignments calibration: (a) Dataset 1; and (b) Dataset 2.

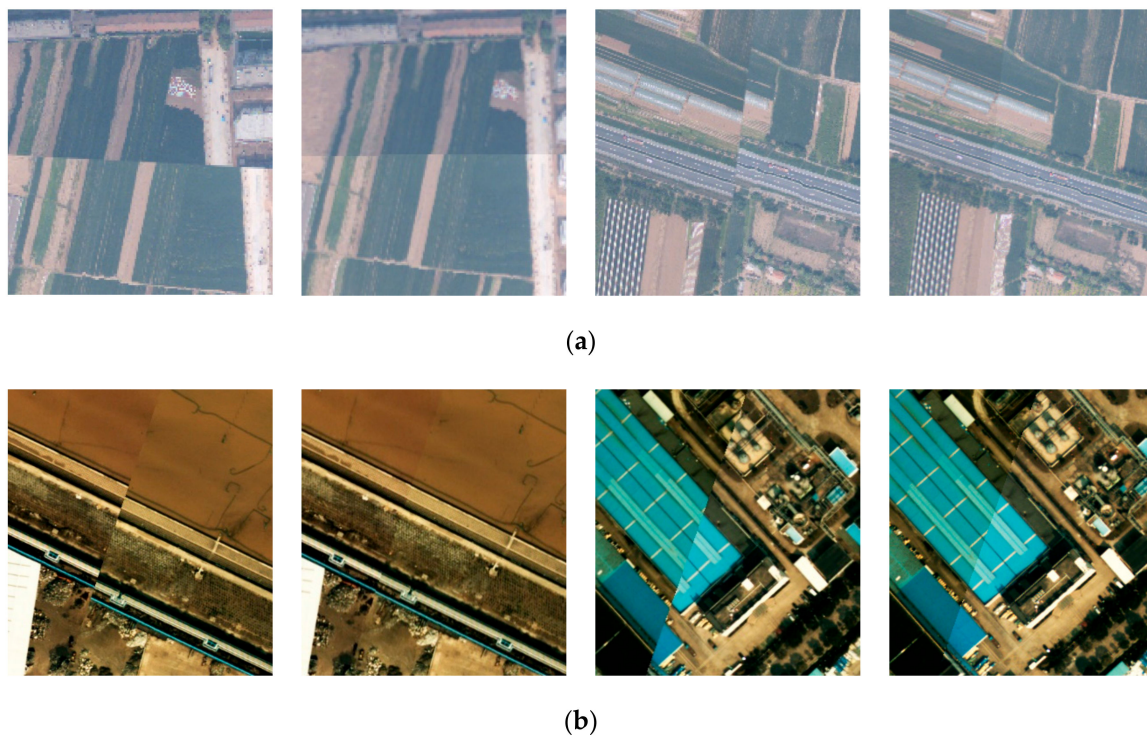


Figure 12. Local area of georeferenced images before and after boresight misalignments calibration: (a) Dataset 1; and (b) Dataset 2.

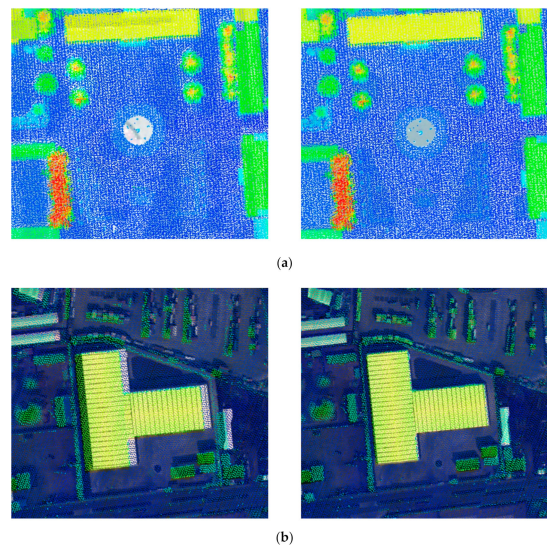


Figure 13. Overlapping LiDAR points with georeferenced images: (a) Dataset 1; and (b) Dataset 2.

4. Discussions

The most important step in boresight misalignments calibration is the establishment of correspondence between tie points and control information. Tie points detection and VCPs selection are two core stages in this study. In the former stage, tie points, which are the basic input data to perform the method, were detected and matched by SURF algorithm, which is a mature feature detection and matching algorithm. Since the image has been positioned using the initial EOE and IOE provided by the POS system, only the feature points in the center S -scale window where the feature points are located are paired. This step can significantly improve the extraction efficiency of tie points. In the later stage, we design an automated tie points selection procedure to perform various assessments to ensure a high quality of point selection from the large numbers of points. These assessments involve

planarity tests to select points located over planar surfaces, slope tolerances calculation to avoid the selected points are located in areas with large vertical differences, and reliability checking to ensure that the selected points have enough redundancy for blunder detection.

The number of VCPs used and the impact on the calibration results were studied as well. Figure 14 shows the accuracy curve of georeferenced images versus the number of VCPs. The curve tends to be horizontal when the number of VCPs is larger than 16. In addition, the distribution of tie points is an equally important factor. An ideal distribution pattern requires that all the tie points be distributed evenly in the whole surveying area.

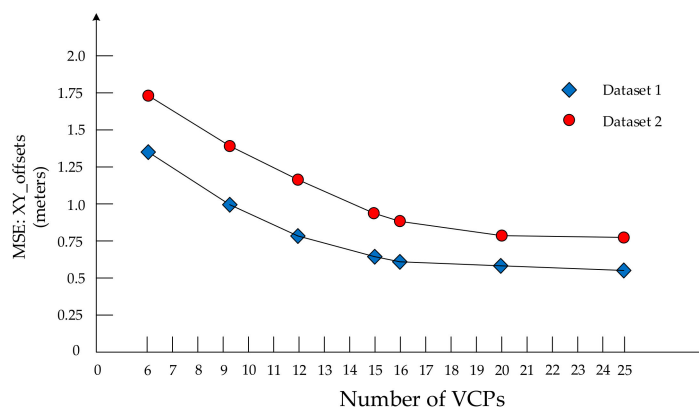


Figure 14. The influence of the number of VCPs.

5. Conclusions

In this paper, a new direct georeferencing method based on boresight misalignments calibration is presented. We validated this method with two different camera systems. Experimental results show that the planar accuracy of the resultant georeferenced images after boresight misalignments calibration is increased significantly compared to those produced directly by using the initial orientation elements from POS data. Moreover, the georeferenced images produced by the proposed method registered the LiDAR points much better. The proposed method allows for boresight misalignments calibration in areas where new calibration fields cannot be established or where rapid response is required in situations such as earthquake relief.

Theoretically, the method is also applicable when the point cloud is not collected simultaneously with the images. Such point cloud can be collected previously by a LiDAR system or generated from other means such as stereo-image matching, only if the images are collected with auxiliary of a POS system. However, such experiments were not conducted in the study and were left for further study. Meanwhile, because the boresight misalignments were contained in the collinear equations as an extra rotation in addition to the rotation caused by the three angular elements of the exterior orientation elements when a virtual control point was projected onto the image space, it is unlikely that other sources of errors can be included as well. Other adjustment models may overcome this shortcoming, but it remains for further studies.

Author Contributions: Conceptualization, H.M. (Haichi Ma) and H.M. (Hongchao Ma); Data curation, L.Z.; Formal analysis, H.M. (Haichi Ma); Funding acquisition, H.M. (Hongchao Ma); Investigation, H.M. (Haichi Ma) and W.L.; Methodology, H.M. (Haichi Ma); Project administration, H.M. (Hongchao Ma); Resources, H.M. (Hongchao Ma); Software, H.M. (Haichi Ma) and K.L.; Supervision, H.M. (Hongchao Ma); Validation, H.M. (Haichi Ma), K.L. and L.Z.; Visualization, H.M. (Haichi Ma); Writing—original draft, H.M. (Haichi Ma); and Writing—review and editing, H.M. (Haichi Ma). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (2018YFB0504500), National Natural Science Foundation of China (Grant numbers 41601504 and 61378078), and National High Resolution Earth Observation Foundation (11-Y20A12-9001-17/18 and 11-H37B02-9001-19/22).

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

References

1. Kobler, A.; Pfeifer, N.; Ogrinc, P.; Todorovski, L.; Oštir, K.; Džeroski, S. Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. *Remote. Sens. Environ.* **2007**, *108*, 9–23. [[CrossRef](#)]
2. Polat, N.; Uysal, M.M. Investigating performance of Airborne LiDAR data filtering algorithms for DTM generation. *Measurement* **2015**, *63*, 61–68. [[CrossRef](#)]
3. Ma, H.; Zhou, W.; Zhang, L.; Wang, S. Decomposition of small-footprint full waveform LiDAR data based on generalized Gaussian model and grouping LM optimization. *Meas. Sci. Technol.* **2017**, *28*, 045203. [[CrossRef](#)]
4. Meng, X.; Wang, L.; Currit, N. Morphology-based Building Detection from Airborne Lidar Data. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 437–442. [[CrossRef](#)]
5. Hamraz, H.; Jacobs, N.; Contreras, M.A.; Clark, C.H. Deep learning for conifer/deciduous classification of airborne LiDAR 3D point clouds representing individual trees. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 219–230. [[CrossRef](#)]
6. Axelsson, P. DEM generation from laser scanner data using adaptive TIN models. *Int. Arch. Photogram. Remote Sens.* **2000**, *33*, 110–118.
7. Mitishita, E.; Cortes, J.; Centeno, J.A.S. Indirect georeferencing of digital SLR imagery using signalised lidar control points. *Photogramm. Rec.* **2011**, *26*, 58–72. [[CrossRef](#)]
8. Huang, Y.; Zhuo, L.; Tao, H.; Shi, Q.; Liu, K. A Novel Building Type Classification Scheme Based on Integrated LiDAR and High-Resolution Images. *Remote Sens.* **2017**, *9*, 679. [[CrossRef](#)]
9. Castagno, J.; Atkins, E.M. Roof Shape Classification from LiDAR and Satellite Image Data Fusion Using Supervised Learning. *Sensors* **2018**, *18*, 3960. [[CrossRef](#)]
10. Luo, S.; Wang, C.; Xi, X.; Zeng, H.; Li, D.; Xia, S.; Wang, P. Fusion of Airborne Discrete-Return LiDAR and Hyperspectral Data for Land Cover Classification. *Remote Sens.* **2015**, *8*, 3. [[CrossRef](#)]
11. Xu, Z.; Guan, K.; Casler, N.; Peng, B.; Wang, S. A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 423–434. [[CrossRef](#)]
12. Han, X.; Wang, H.; Lu, J.; Zhao, C. Road detection based on the fusion of Lidar and image data. *Int. J. Adv. Robot. Syst.* **2017**, *14*, 1–10. [[CrossRef](#)]
13. Sameen, M.; Pradhan, B. A Two-Stage Optimization Strategy for Fuzzy Object-Based Analysis Using Airborne LiDAR and High-Resolution Orthophotos for Urban Road Extraction. *J. Sens.* **2017**, *2017*, 1–17. [[CrossRef](#)]
14. Matsuki, T.; Yokoya, N.; Iwasaki, A. Hyperspectral Tree Species Classification of Japanese Complex Mixed Forest With the Aid of Lidar Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2177–2187. [[CrossRef](#)]
15. Pham, L.; Brabyn, L.; Ashraf, S. Combining QuickBird, LiDAR, and GIS topography indices to identify a single native tree species in a complex landscape using an object-based classification approach. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 187–197. [[CrossRef](#)]
16. Gneeniss, A.; Mills, J.P.; Miller, P.E. In-flight photogrammetric camera calibration and validation via complementary lidar. *ISPRS J. Photogramm. Remote Sens.* **2015**, *100*, 3–13. [[CrossRef](#)]
17. Grussenmeyer, P.; Al Khalil, O. Solutions for Exterior Orientation in Photogrammetry: A Review. *Photogramm. Rec.* **2002**, *17*, 615–634. [[CrossRef](#)]
18. Mikolajczyk, K.; Schmid, C. Indexing Based on Scale Invariant Interest Points. In Proceedings of the Eighth IEEE International Conference on Computer Vision ICCV, Vancouver, BC, Canada, 7–14 July 2001; Volume 1, pp. 525–531. [[CrossRef](#)]
19. Honkavaara, E.; Markelin, L.; Ahokas, E.; Kuittinen, R.; Peltoniemi, J. Calibrating digital photogrammetric airborne imaging systems in a test field. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 555–560.
20. Yastikli, N.; Jacobsen, K. Direct Sensor Orientation for Large Scale Mapping—Potential, Problems, Solutions. *Photogramm. Rec.* **2005**, *20*, 274–284. [[CrossRef](#)]
21. Gautam, D.; Lucieer, A.; Watson, C.; McCoull, C. Lever-arm and boresight correction, and field of view determination of a spectroradiometer mounted on an unmanned aircraft system. *ISPRS J. Photogramm. Remote Sens.* **2019**, *155*, 25–36. [[CrossRef](#)]
22. Seo, J.; Lee, H.K.; Jee, G.; Park, C.G. Lever arm compensation for GPS/INS/Odometer integrated system. *Int. J. Control Autom. Syst.* **2006**, *4*, 247–254.

23. Heier, H.; Kiefner, M.; Zeitler, W. Calibration of the Digital Modular Camera. In Proceedings of the FIG XXII International Congress, Washington, DC, USA, 19–26 April 2002; p. 11.
24. Skaloud, J.; Cramer, M.; Schwarz, K.P. Exterior orientation by direct measurement of position and attitude. *Bmc Health Serv. Res.* **1996**, *8*, 789.
25. Skaloud, J. Optimizing Georeferencing of Airborne Survey Systems by INS/GPS. Ph.D. Thesis, Department of Geomatics Engineering, The University of Calgary, Calgary, Alberta, 1999.
26. Heipke, C.; Jacobsen, K.; Wegmann, H. The OEEPE Test on Integrated Sensor Orientation—Results of Phase I. In Proceedings of the Photogrammetric Week, Stuttgart, Germany, 24–28 September 2001; Fritsch, D., Spiller, R., Eds.; Wichmann Verlag: Berlin, Germany, 2001; pp. 195–204.
27. Jacobsen, K. Aspects of Handling Image Orientation by Direct Sensor Orientation. In Proceedings of the American Society of Photogrammetry and Remote Sensing Annual Convention, St. Louis, MO, USA, 23–27 April 2001; pp. 629–633.
28. Mostafa, M.M.R.; Corporation, A. Camera/IMU Bore-sight Calibration: New Advances and Performance Analysis. In Proceedings of the American Society of Photogrammetry and Remote Sensing Annual Conference, Washington, DC, USA, 21–26 April 2002; p. 12.
29. Honkavaara, E. Calibration field structures for GPS/IMU/camera-system calibration. *Photogramm. J. Finl.* **2003**, *18*, 3–15.
30. Honkavaara, E. Calibration in direct georeferencing: Theoretical considerations and practical results. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 1207–1208.
31. Jacobsen, K. Direct integrated sensor orientation—Pros and Cons. *Int. Arch. Photogramm. Remote Sens.* **2004**, *35*, 829–835.
32. Filho, L.E.; Mitishita, E.A.; Kersting, A.P.B. Geometric Calibration of an Aerial Multihead Camera System for Direct Georeferencing Applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1–12. [[CrossRef](#)]
33. Kruck, E.J. Simultaneous calibration of digital aerial survey cameras. In Proceedings of the EuroSDR Commission I and ISPRS Working Group 1/3 Workshop EuroCOW, Castelldefels, Spain, 25–27 January 2006; p. 7.
34. Jacobsen, K. Geometric handling of large size digital airborne frame camera images. *Opt. 3D Meas. Tech.* **2007**, *8*, 164–171.
35. Delara, R.; Mitishita, E.; Habib, A. Bundle adjustment of images from nonmetric CCD camera using LiDAR data as control points. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *20*, 13–18.
36. Habib, A.; Ghanma, M.; Mitishita, E.; Kim, E.; Kim, C.J. Image georeferencing using LIDAR data. In Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05, Seoul, Korea, 29 July 2005; Volume 2, pp. 1158–1161. [[CrossRef](#)]
37. Habib, A.; Ghanma, M.; Morgan, M.; Al-Ruzouq, R. Photogrammetric and Lidar Data Registration Using Linear Features. *Photogramm. Eng. Remote. Sens.* **2005**, *71*, 699–707. [[CrossRef](#)]
38. Kwak, T.-S.; Kim, Y.-I.; Yu, K.-Y.; Lee, B.K. Registration of aerial imagery and aerial LiDAR data using centroids of plane roof surfaces as control information. *KSCE J. Civ. Eng.* **2006**, *10*, 365–370. [[CrossRef](#)]
39. Liu, X.; Zhang, Z.; Peterson, J.; Chandra, S. Lidar-derived high quality ground control information and DEM for image orthorectification. *Geoinformatica* **2007**, *11*, 37–53. [[CrossRef](#)]
40. Yastikli, N.; Toth, C.; Brzezinska, D. In-Situ camera and boresight calibration with lidar data. In Proceedings of the 5th International Symposium on Mobile Mapping Technology, Padua, Itália, 29–31 May 2007; p. 6.
41. Mitishita, E.; Habib, A.F.; Centeno, J.A.S.; Machado, A.M.L.; Lay, J.C.; Wong, C. Photogrammetric and lidar data integration using the centroid of a rectangular roof as a control point. *Photogramm. Rec.* **2008**, *23*, 19–35. [[CrossRef](#)]
42. Ding, M.; Lyngbaek, K.; Zakhor, A. Automatic registration of aerial imagery with untextured 3D LiDAR models. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 24–26 January 2008; pp. 23–28. [[CrossRef](#)]
43. Wang, L.; Neumann, U. A Robust Approach for Automatic Registration of Aerial Images with Untextured Aerial LiDAR Data. In Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 2623–2630. [[CrossRef](#)]
44. Wildan, F.; Aldino, R.; Aji, P.P. Application of LIDAR Technology for GCP Determination in Papua Topographic Mapping Scale 1:50,000. In Proceedings of the 10th Annual Asian Conference & Exhibition on Geospatial Information, Technology & Applications, Jakarta, Indonesia, 17–19 October 2011; p. 13.

45. Siying, C.; Hongchao, M.; Yinchao, Z.; Liang, Z.; Jixian, X.; He, C. Bore-sight Calibration of Airborne LiDAR System Without Ground Control Points. *IEEE Geosci. Remote Sens. Lett.* **2011**, *9*, 85–89. [[CrossRef](#)]
46. Gneeniss, A.S. Integration of LiDAR and Photogrammetric Data for Enhanced Aerial Triangulation and Camera Calibration. Ph.D. Thesis, Newcastle University, Newcastle upon Tyne, UK, 2014.
47. Rusu, R.B.; Cousins, S. 3D Is here: Point Cloud Library (PCL). In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1–4. [[CrossRef](#)]
48. Dharmapuri, S.S. Vertical accuracy validation of LiDAR data. *LiDAR Mag.* **2014**, *4*, 1.
49. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [[CrossRef](#)]
50. Goshtasby, A.; Stockman, G.; Page, C. A Region-Based Approach to Digital Image Registration with Subpixel Accuracy. *IEEE Trans. Geosci. Remote Sens.* **1986**, *24*, 390–399. [[CrossRef](#)]
51. You, J.; Bhattacharya, P. A wavelet-based coarse-to-fine image matching scheme in a parallel virtual machine environment. *IEEE Trans. Image Process.* **2000**, *9*, 1547–1559. [[CrossRef](#)]
52. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
53. Gonçalves, H.; Corte-Real, L.; Gonçalves, J.A. Automatic Image Registration Through Image Segmentation and SIFT. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2589–2600. [[CrossRef](#)]
54. Ye, F.; Su, Y.; Xiao, H.; Zhao, X.; Min, W. Remote Sensing Image Registration Using Convolutional Neural Network Features. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 232–236. [[CrossRef](#)]
55. Jeong, J. Imaging Geometry and Positioning Accuracy of Dual Satellite Stereo Images: A Review. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 235–242. [[CrossRef](#)]
56. Kodors, S.; Kangro, I. Simple method of LiDAR point density definition for automatic building recognition. *Eng. Rural Dev.* **2016**, *5*, 415–424.
57. Jacobsen, K. *BLUH Bundle Block Adjustment*; Program User Manual; Institute of Photogrammetry and GeoInformation, Leibniz University Hannover: Hannover, Germany, 2008; p. 34.
58. Baumker, M.; Heimes, F.J. New Calibration and Computing Method for Direct Georeferencing of Image and Scanner Data Using the Position and Angular Data of an Hybrid Inertial Navigation System. *Integr. Sens. Orientat.* **2002**, *43*, 197–212.
59. Zhao, H.; Zhang, B.; Wu, C.; Zuo, Z.; Chen, Z. Development of a Coordinate Transformation method for direct georeferencing in map projection frames. *ISPRS J. Photogramm. Remote Sens.* **2013**, *77*, 94–103. [[CrossRef](#)]
60. Mistry, D.; Banerjee, A. Comparison of Feature Detection and Matching Approaches: SIFT and SURF. *GRD J. Global Res. Dev. J. Eng.* **2017**, *2*, 7–13.



© 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/>).