

# ACCURATE REGISTRATION OF AERIAL IMAGES AND ALS-POINTCLOUD VIA AUTOMATED JUNCTION MATCHING AND PLANAR CONSTRAINTS

Yi Wan, Yongjun Zhang\*, Guangshuai Wang, Xinyi Liu

School of Remote Sensing and Information Engineering, Wuhan University, 430079 Wuhan Hubei Province, China -  
(yi.wan, zhangyj, gs\_wang, liuxy0319)@whu.edu.cn

Commission VI, WG VI/4

**KEY WORDS:** Photogrammetry, Registration, Image Matching, Junction, LiDAR

## ABSTRACT:

Accurate geometric registration of images and pointclouds is the key step of many 3D-reconstruction or 3D-sensing tasks. In this paper, a novel L-junction based approach is proposed for semi-automatic accurate registration of aerial images and the airborne laser scanning (ALS) point-cloud in urban areas. The approach achieves accurate registration by associating the LiDAR points with the local planes extracted via L-junction detection and matching from multi-view aerial images. An L-junction is an intersection of two line-segments. Through the forward intersection of multi-view corresponding L-junctions, an accurate local junction-plane can be obtained. In the proposed approach, L-junction is manually collected from one view on the flat object-surfaces like walls, roads, and roofs and then automatically matched to other views with the aid of epipolar-geometry and vanishing-point constraints. Then, a plane-constrained bundle block adjustment of the image-orientation parameters is conducted, where the LiDAR points are treated as reference data. The proposed approach was tested with two datasets collected in Guangzhou city and Ningbo city of China. The experimental results showed that the proposed approach had better accuracy than the closest-point based method. The horizontal/vertical registration RMS of the proposed approach reached 4.21cm/5.72cm in Guangzhou dataset and 4.46cm/4.34cm in Ningbo dataset, which was much less than the average LiDAR-point distance (over 25cm in both datasets) and was very close to the image GSDs (3.2cm in Guangzhou and 4.8cm in Ningbo) and the a-priori ranging accuracy of the ALS equipment (about 3cm).

## 1. INTRODUCTION

The recent two decades have witnessed the rapid development of the digital cameras and the Light Detection and Ranging (LiDAR) sensors. With the acquisition expense gets lower, the integration of both data sources (images and LiDAR pointclouds) becomes popular in dealing with 3D-sensing tasks like topographic mapping, auto-pilot, indoor Simultaneous Localization and Mapping (SLAM), 3D reconstruction, 3D object detection (Huang, et al., 2019), etc. For whatever tasks, the accurate geometric registration of the two data sources should be ensured at the beginning (Mishra and Zhang, 2012; Zhang and Lin, 2016). The geometric registration of both data sources is mainly implemented through two types of methods, i.e., the 2D multi-modal image-matching methods and the 3D-feature-registration methods.

In the methods based on multi-modal image-matching, the 3D information in LiDAR data should be projected onto the 2D image-space at the beginning. In (Wong and Orchard, 2008), the intensity of the LiDAR reflections was projected onto the image-space to form the intensity map, and then the registration was achieved by extracting point-matches between the intensity map and the aerial images. In (Shorter and Kasparis, 2008), the edges of the buildings were extracted from the LiDAR pointclouds and projected onto the image-space to be matched with the aerial images through phase correlation. In (Barsai, et al., 2017), the edges of the buildings were extracted from both the LiDAR pointclouds and the aerial images, then the registration was achieved by minimizing the overall distances of the two sets of 2D-edges without establishing explicit correspondences. The above-mentioned methods transform the 3D alignment problem into a 2D multi-modal image-matching problem or a

2D structure-registration problem, and as a result, the alignment accuracy is strongly related to the average LiDAR-point distance and the ground sampling distance (GSD) of aerial images.

In the methods based on 3D feature registration, structures like points, line-segments, or planes are extracted from both the images and the LiDAR point-clouds. The structures extracted from the images need to be back-projected to the 3D-space through the forward-intersection process before the 3D-registration is processed. For the aerial images and the airborne laser scanning (ALS) data acquired over urban areas, the widely existing corners, lines, and planes should be used to get better registration accuracy. Because the airborne laser scanning (ALS) data usually has better ranging accuracy (2cm-5cm) than its ground sample distance (GSD, 20cm-30cm). Extracting structures from LiDAR pointcloud usually begins with extracting planes. Two planes are needed to get a line-segment and three are needed to get a corner point. Line-segments and corner points can be easily extracted from the images, while planes cannot. Although a plane can be determined by three or more tie-points extracted from multi-view images, it is not easy to judge whether this plane is on an existed flat object surface. As a result, many researchers achieved accurate 3D registration through corners and line-segments. In (Gruen and Akca, 2005; Akca, 2007), the point-cloud extracted through dense-match of the aerial images were used for the registration with the LiDAR point-cloud. In (Zheng, et al., 2013; Huang, et al., 2018), the 3D points extracted from the images were constrained with the planes extracted from the LiDAR data through normal-shooting strategy. In (Costa, et al., 2017; Costa and Mitishita, 2019; Peng and Zhang, 2019), the line-segments or the corner points of the

\* Corresponding author

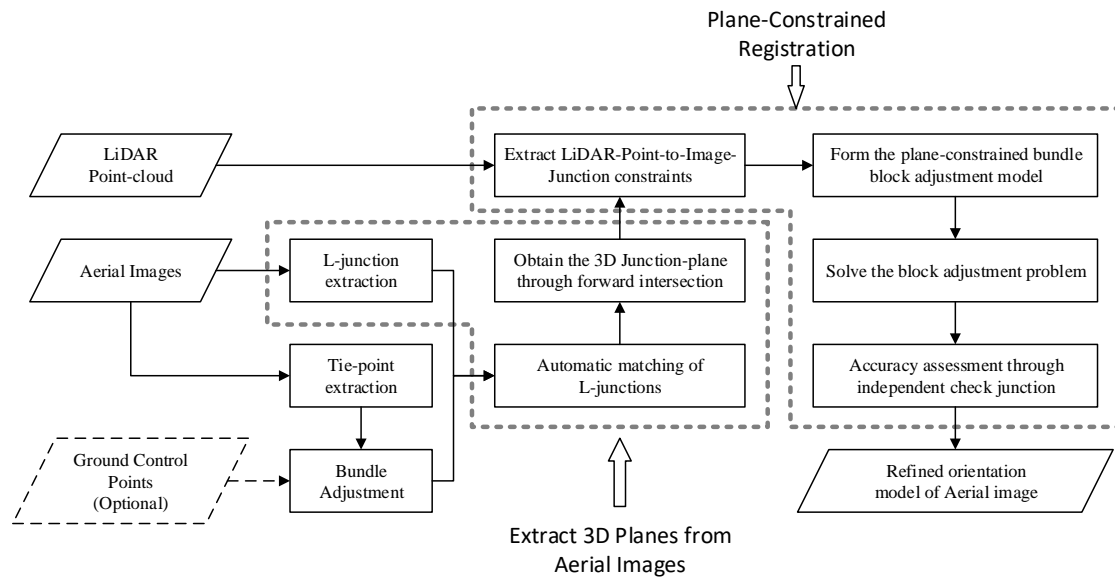


Figure 1. Workflow of the proposed approach.

buildings, extracted from the LiDAR data through intersecting the point-cloud planes, were registered with those extracted from the images. In (Armenakis, et al., 2013), the author discussed the potential of directly using the planes extracted from both the images and the LiDAR data as the constraints for 3D registration. The above-mentioned 3D-registration methods utilize different geometric structures to deal with different tasks, scales, and targets.

In this paper, a new plane-based approach is proposed for the geometric registration of the aerial images and the LiDAR point-clouds in urban areas. The new approach extracts planes from the aerial images through obtaining the L-junctions from multi-view aerial images and then achieves accurate 3D registration through associating the junction-planes with the nearby LiDAR points. The novelty of the approach is in two aspects:

1. A new method of automatically matching the L-junction of the line-segments is proposed, with the aid of epipolar-geometry and the vanishing-point constraints.
2. A new plane-constrained bundle block adjustment model of the aerial images is designed to achieve the accurate registration of the aerial images and the ALS pointcloud.

The workflow of the proposed approach is illustrated in Figure 1. The process has three main steps. First, the initial orientation models of the aerial images are prepared through tie-point matching and POS-aided bundle adjustment. Second, the L-junctions are manually collected from the aerial images and then automatically matched to other views. The collected L-junctions should correspond to real-existed planar object-surfaces. Thus, the 3D planes can be extracted through the forward intersection process of the multi-view L-junctions, and by associate the planes with the nearby LiDAR points, the planar constraints are established. Finally, a plane-constrained bundle block adjustment is operated to the image orientation parameters to achieve accurate 3D registration between the two data sources.

The reminder of this paper is as follows. In Section 2, the method and the related theory of automatic L-junction matching and is introduced. In Section 3, the way of linking the LiDAR-

points and the junction-planes is described and the planar constrained block adjustment model is introduced in detail. In Section 4, two datasets are used to demonstrate the proposed approach. In Section 5, conclusions are drawn.

## 2. EXTRACT PLANES FROM AERIAL IMAGES

As in Figure 1, three steps are needed for extracting planes from aerial images, i.e., collect the L-junctions manually, match the L-junctions automatically, and obtain the 3D junction-plane via forward intersection of the multi-view L-junctions.

### 2.1 Detect L-Junctions

Although there are many well-designed methods for automatically detecting the corners, line-segments, or junction-like structures, this approach chooses to get the first image-correspondence of an L-junction manually and then find its correspondences from other views automatically. Because currently, no automatic detection methods can ensure its results satisfying the following rules.

#### Rules of L-junction Collection:

1. The plane yielded by the two branches of an L-junction must correspond to a real existed flat object-surface.
2. Each of the branch of an L-junction should be either horizontal or vertical.

The first rule ensures that the collected L-junctions and their correspondences can be used to extract real existed planes for 3D-registration (see Figure 2), while the second rules ensures that the L-junctions are easy to match under the vanishing-point constraints. The images-space L-junctions is a two-branch junction, which is denoted as:

$$J = \{c, \vec{p}, \vec{q}\} \quad (1)$$

where  $c$  is the junction-centre,  $\vec{p}$  and  $\vec{q}$  are two branches. The branch of a L-junction is a line-segment from the junction-centre to its terminal point (denoted as  $p$  and  $q$ ).



Figure 2. The L-junctions in white solid lines can be used to extract planes, while those in red dash lines cannot.

## 2.2 Match the L-junctions

It is quite easy for a person to find an L-junction from an aerial image that satisfies the Rules of L-junction Collection in Section 2.1. However, it is quite difficult to find correspondences for an L-junction from unordered aerial images. Even the images have been bundle-adjusted, a well-designed graphical user-interface is needed to manually match them and the job is still very time-costing because one L-junction may have tens of image-correspondences in a set of oblique aerial images. Thus, an automatic method is proposed.

In this approach, the purpose of matching the L-junctions is to extract the 3D-planes yielded by the junction-branches. Thus, two image-space L-junctions ( $J = \{c, \bar{p}, \bar{q}\}$  and  $J' = \{c', \bar{p}', \bar{q}'\}$ ) will be treated as correspondences when:

1. The junction-centre  $c$  and  $c'$  are corresponding points;
2. The junction-branch  $\bar{p}$  and  $\bar{p}'$  are on two corresponding straight-lines;
3. The junction-branch  $\bar{q}$  and  $\bar{q}'$  are on two corresponding straight-lines;

As is described above, the terminate points of the corresponding branches do not need to be corresponding points. During the automatic matching process, the correspondence (denoted as  $J'$ ) of the junction  $J$  can be searched under the following constraints (see Figure 3).

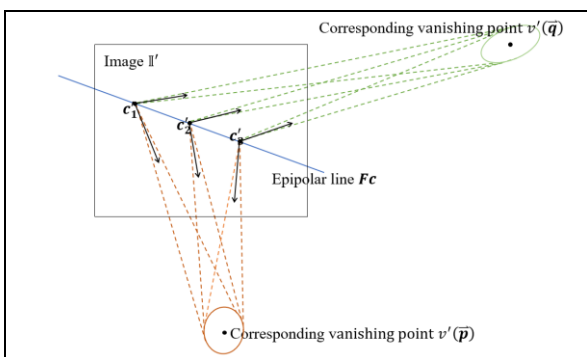


Figure 3. Searching the corresponding L-junctions under epipolar-geometry and constraints of two vanishing-points.

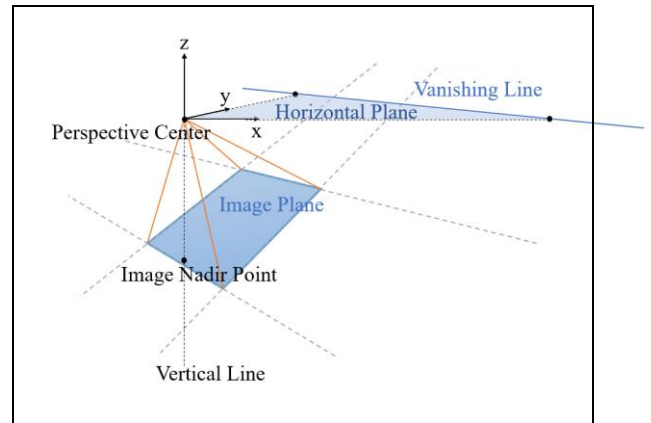


Figure 4. The vanishing line and the image-space nadir point.

### Constraints for L-junction Matching:

1. The junction-centre  $c$  and  $c'$  should follow the principle of epipolar-geometry.
2. The straight-lines on which the corresponding junction-branches lies should pass the corresponding vanishing points.

As a result, the epipolar-geometry and the constraints of two vanishing points can be used to aid the matching of L-junctions. Since the orientation model of the aerial images can be treated as an apriori in the proposed approach, the epipolar-geometric relationship of any image-pairs can be easily deduced. Denote the fundamental matrix between image  $I$  and  $I'$  as  $F$ , the epipolar-geometric-based constraint is:

$$\text{dist}(c', Fc) < \mu_i \quad (2)$$

where  $\text{dist}(\cdot)$  means the distance between the point  $c'$  and the epipolar-line  $Fc$ ;  $\mu_i$  is the threshold of the distance.

In order to match the L-junctions with the branch-direction constrained, we need to firstly locate the vanishing points for the branches of the artificially collected L-junctions, and then find the correspondences of the vanishing points from the other views. Vanishing point is the projection of a 3D infinite point. According to the Rules of Artificial L-junction Collection in Section 2.1, the branch of an L-junction is either horizontal or vertical. Thus, the vanishing point of a branch should be either the image-space nadir point or the intersection of the branch (the straight line that contains the branch) and the vanishing line (see Figure 4). Denote  $l(\bar{p})$  as the straight-line that contains the branch  $\bar{p}$  and  $v(\bar{p})$  as the vanishing point corresponding to  $l(\bar{p})$ . Denote the ground correspondence of  $v(\bar{p})$  as  $\mathcal{V}(\bar{p})$ . If  $v(\bar{p})$  is the nadir point, then the homogeneous coordinate of  $\mathcal{V}(\bar{p})$  is  $(0 \ 0 \ -1 \ 0)^T$ . If  $v(\bar{p})$  is not the nadir point, then the homogeneous coordinate of  $\mathcal{V}(\bar{p})$  is  $(\cos\theta \ \sin\theta \ 0 \ 0)^T$ , which has only one degree of freedom and can be easily solved with the colinear equation:

$$v(\bar{p}) = [R | T] \cdot \mathcal{V}(\bar{p}) \quad (3)$$

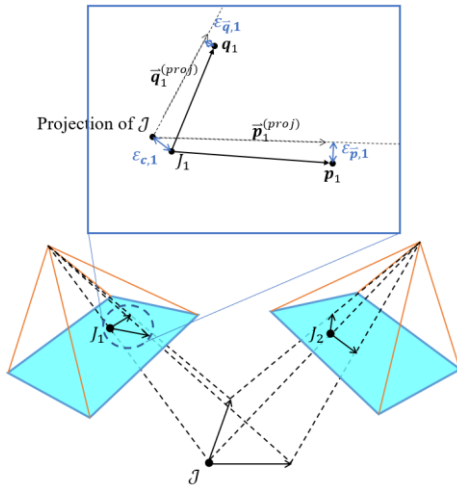


Figure 5. The forward intersection of L-junctions.

Using similar steps, we can get  $v(\bar{\mathbf{q}})$  and  $v(\bar{\mathbf{q}})$ . Finally, the image-space correspondences of  $v(\bar{\mathbf{p}})$  and  $v(\bar{\mathbf{q}})$  can be obtained by projecting  $v(\bar{\mathbf{p}})$  and  $v(\bar{\mathbf{q}})$  onto the other images:

$$\begin{aligned} v'(\bar{\mathbf{p}}) &= [\mathbf{R}' | \mathbf{T}'] \cdot v(\bar{\mathbf{p}}) \\ v'(\bar{\mathbf{q}}) &= [\mathbf{R}' | \mathbf{T}'] \cdot v(\bar{\mathbf{q}}) \end{aligned} \quad (4)$$

With the vanishing point  $v'(\bar{\mathbf{p}})$  and  $v'(\bar{\mathbf{q}})$ , the vanishing-point constraints are:

$$\text{angle}(\overline{v'(\bar{\mathbf{p}})}, \bar{\mathbf{p}}) < \mu_2 \text{ and } \text{angle}(\overline{v'(\bar{\mathbf{q}})}, \bar{\mathbf{q}}) < \mu_2 \quad (5)$$

where  $\text{angle}(\cdot)$  means the angle between two lines;  $\mu_2$  is the angular threshold.

The proposed automatic matching method has three steps:

1. Locate the vanishing points  $v(\bar{\mathbf{p}})$  and  $v(\bar{\mathbf{q}})$  on the image  $\mathbf{I}$ , and deduce the vanishing points  $v'(\bar{\mathbf{p}})$  and  $v'(\bar{\mathbf{q}})$  on the image  $\mathbf{I}'$ .
2. Automatically detect line-segments on image  $\mathbf{I}'$ ;
3. Form the L-junctions with the line-segments on image  $\mathbf{I}'$  and select the one that satisfies all the constraints in (2) and (5).

If more than one L-junctions satisfy the constraints, manually check is needed to find the correct one.

### 2.3 Forward Intersection and Plane Extraction

An object-space 3D L-junction is composed by a 3D point and two associating directions. Thus, at least seven parameters are needed to express an 3D L-junction, which is denoted as:

$$\mathcal{J} = \{X, Y, Z, \phi_p, \theta_p, \phi_q, \theta_q\} \quad (6)$$

where  $X$ ,  $Y$ , and  $Z$  are the coordinates of the centre point;  $\phi_p$  and  $\theta_p$  express the direction of the branch  $\bar{P}$ ;  $\phi_q$  and  $\theta_q$  express the direction of the branch  $\bar{Q}$ . The forward intersection

of L-junctions has two steps, i.e., obtaining the initial value  $\mathcal{J}_0$  and then optimize it through least-square-based multi-view forward intersection. The initial coordinate of the junction-centre can be obtained by direct forward intersection with any two of the image-space junction-centres, while the initial branch-direction can be obtained from the 3D infinite point  $\mathcal{V}(\bar{\mathbf{p}})$  and  $\mathcal{V}(\bar{\mathbf{q}})$  used in L-junction matching.

The projection of the 3D L-junction  $\mathcal{J}$  is in the following steps. First, project the junction-centre  $(X, Y, Z)$  onto the image-space. Then, project an arbitrary 3D point on each of the 3D-branch onto the image-space and thus we can get the 2D branches. Denote the projected junction-centre on the  $j$ -th image as  $\mathbf{c}_j^{(proj)}$ , and the two projected junction-branch as  $\bar{\mathbf{p}}_j^{-(proj)}$  and  $\bar{\mathbf{q}}_j^{-(proj)}$ . The problem of optimizing the 3D L-junction  $\mathcal{J}$  is denoted as:

$$\begin{aligned} \mathcal{J} &\leftarrow \arg \min \boldsymbol{\varepsilon}_{\mathcal{J}}^T \boldsymbol{\varepsilon}_{\mathcal{J}} \\ \boldsymbol{\varepsilon}_{\mathcal{J}}^T &= [\dots \quad \varepsilon_{c,j} \quad \varepsilon_{\bar{\mathbf{p}},j} \quad \varepsilon_{\bar{\mathbf{q}},j} \quad \dots] \\ \varepsilon_{c,j} &= \text{dist}(\mathbf{c}_j, \mathbf{c}_j^{(proj)}) \\ \varepsilon_{\bar{\mathbf{p}},j} &= \text{dist}(\bar{\mathbf{p}}_j, \bar{\mathbf{p}}_j^{-(proj)}) \\ \varepsilon_{\bar{\mathbf{q}},j} &= \text{dist}(\bar{\mathbf{q}}_j, \bar{\mathbf{q}}_j^{-(proj)}) \end{aligned} \quad (7)$$

In the optimization problems,  $J_j = \{\mathbf{c}_j, \bar{\mathbf{p}}_j, \bar{\mathbf{q}}_j\}$  is an artificially collected or an automatically matched 2D junction on the  $j$ -th image. The cost value  $\varepsilon_{c,j}$  is a point-to-point distance, while  $\varepsilon_{\bar{\mathbf{p}},j}$  and  $\varepsilon_{\bar{\mathbf{q}},j}$  are point-to-line distances (see Figure 5). With the optimized 3D junction parameters, the 3D plane can be easily obtained by the junction centre and two junction-branches. Denote the homogeneous coordinate of the 3D plane yielded by  $\mathcal{J}$  as  $\mathbf{P}(\mathcal{J})$ . The 3D alignment between the aerial images and the LiDAR points will be operated by forcing the LiDAR points near  $\mathbf{P}(\mathcal{J})$  to be on it.

## 3. PLANE CONSTRAINED REGISTRATION

There are three steps to operate the plane-constrained 3D registration. First, the LiDAR points near the 3D junction-plane are searched to form the planar constraints. Second, a bundle block adjustment model with additional point-to-plane cost functions is formed and solved to refine the orientation model of the aerial images, so as to achieve the accurate registration. Finally, the registration accuracy is assessed by the independent check junctions (ICJs).

### 3.1 Point-to-Plane Cost Functions

The parameters of the 3D junction  $\mathcal{J}$  do not involve the length of the branches because they will not be used in the forward intersection process and the following block adjustment. However, estimating the branch-length can help us find the related LiDAR points. With the junction  $\mathcal{J}$  and two branch-lengths  $L(\bar{\mathbf{p}})$  and  $L(\bar{\mathbf{q}})$ , the LiDAR points within the 3D cube formed by the junction-plane and a thickness threshold  $\mu_H$  are related with the junction  $\mathcal{J}$  to build the point-to-plane constraint (see Figure 6).

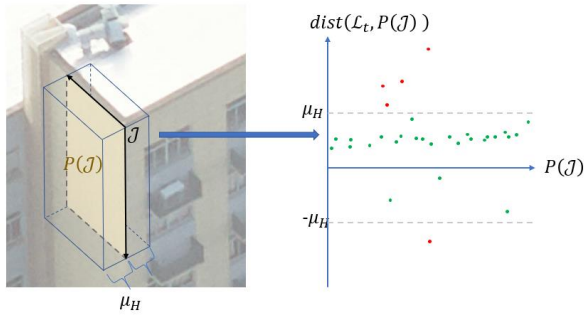


Figure 6. The points (green) within the buffer of the junction-plane will be constrained to the L-junction.

The point-to-plane constraint can be denoted as the following cost function with the  $t$ -th LiDAR point  $\mathcal{L}_t$  and the  $i$ -th 3D L-junction  $\mathcal{J}_i$ ,

$$\varepsilon_{\mathcal{J}_i, \mathcal{L}_t} = \text{dist}(\mathcal{L}_t, \mathbf{P}(\mathcal{J}_i)), \text{ with weight } W_{i,t} \quad (8)$$

where  $\text{dist}(\mathcal{L}_t, \mathbf{P}(\mathcal{J}_i))$  is the point-to-plane distance; the weight value  $W_{i,t}$  will be fixed as the one at the first iteration of the block adjustment, and in the latter iterations will be modified according to the point-to-plane distance calculated with the updated image orientation model:

$$W_{i,t} \leftarrow \begin{cases} 1, & \text{if } \text{dist}(\mathcal{L}_t, \mathbf{P}^*(\mathcal{J}_i)) \leq \mu_p \\ \mu_p^2 / \text{dist}^2(\mathcal{L}_t, \mathbf{P}^*(\mathcal{J}_i)), & \text{if } \text{dist}(\mathcal{L}_t, \mathbf{P}^*(\mathcal{J}_i)) > \mu_p \end{cases} \quad (9)$$

where  $\mathbf{P}^*(\mathcal{J}_i)$  is the junction-plane of  $\mathcal{J}_i$  calculated by the orientation parameters updated in the last iteration.

### 3.2 Plane-constrained Bundle Block Adjustment

Combining the cost function of L-junction projection, tie-point projection, and the point-to-plane constraints, a mixed bundle block adjustment model is denoted as:

$$\begin{aligned} \mathbf{x} &\leftarrow \arg \min \boldsymbol{\varepsilon}^T \mathbf{P} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}^T &= [\boldsymbol{\varepsilon}_{\mathcal{J}_i}^T \quad \cdots \quad \varepsilon_{\mathcal{J}_i, \mathcal{L}_t} \quad \cdots \quad \boldsymbol{\varepsilon}_{\mathcal{P}_i}^T \quad \cdots] \\ \mathbf{x}^T &= [\mathbf{t}_j \quad \cdots \quad \mathcal{P}_i \quad \cdots \quad \mathcal{J}_i \quad \cdots \quad \mathcal{C}] \end{aligned} \quad (10)$$

where  $\boldsymbol{\varepsilon}$  denotes the cost functions;  $\mathbf{x}$  denotes the variables to be optimized;  $\mathbf{t}_j$  denotes the exterior orientation parameters of the  $j$ -th image;  $\mathcal{P}_i$  is the coordinate of the  $i$ -th tie-point;  $\mathcal{J}_i$  denotes the parameters of the  $i$ -th L-junction;  $\mathcal{C}$  denotes the interior orientation parameters and the distortion parameters of the aerial camera;  $\boldsymbol{\varepsilon}_{\mathcal{J}_i}^T$  denotes the projection-cost of the  $i$ -th L-junction as in (7);  $\varepsilon_{\mathcal{J}_i, \mathcal{L}_t}$  denotes the point-to-plane costs as in (8);  $\boldsymbol{\varepsilon}_{\mathcal{P}_i}^T$  denotes the projection-cost of the  $i$ -th tie-point. Compared with the conventional bundle adjustment, the model in (10) does not involve the cost functions for ground control points and POS data, but involves the cost function for L-junction projection and for plane constraints.

Since the aerial images have been pre-bundle-adjusted before their 3D registration with the LiDAR points, the pre-adjustment results (tie-point locations and image orientation parameters) can be used as initials. The L-junctions can be treated as control points that however can only control one direction (the normal direction of the junction-plane) of the 3D space. Thus, the block adjustment problem (10) can be solved by iterative linearization and variable update as we do for conventional bundle adjustment. During the iteration, the weight values of the (junction- and tie-point-) projection costs are fixed, while the weight values of the point-plane-constraining costs are updated according to (9). With the L-junctions carefully selected according to the Rules of Artificial L-junction Collection, the solving process can converge with several iterations.

### 3.3 Plane-based Accuracy Assessment

The L-junctions not used as control junctions (CJs) can be used as ICJs. In the accuracy assessment, the RMS of the distances between a (horizontal or vertical) junction-plane and the associated LiDAR points can be used to evaluate the (horizontal or vertical) registration error near this L-junction.

## 4. EXPERIMENTS AND ANALYSIS

In this section, the proposed approach was tested with two datasets acquired over Guangzhou and Ningbo of China (see Figure 7). The detailed information of the aerial images and the LiDAR data is given in Table 1. Both the LiDAR point-clouds have been geometrically rectified before the test and had better than 10cm RMSE. The accuracy of the Ningbo LiDAR data has been evaluated with the standard 1:500 topographic maps. The results showed that the horizontal RMSE was about 6.9cm and the vertical RMS was about 7.9cm. The ranging error of the LiDAR sensor was about 3cm. The images and the LiDAR data were acquired at different time.

| Location      |                       | Guangzhou                  | Ningbo                       |
|---------------|-----------------------|----------------------------|------------------------------|
| Aerial Images | Flying Height         | 500m                       | 900m                         |
|               | Spatial Res.          | 3.2cm                      | 4.8cm                        |
|               | Image Size            | 10336×7788                 | 10608×8708                   |
|               | #Cameras              | 5                          | 5                            |
|               | #Images               | 2415                       | 1541                         |
| LiDAR Points  | Mean Point Dist.      | 25cm (hor.)<br>50cm (ver.) | 30cm (hor.)<br>90cm (ver.)   |
|               | RMS of Point accuracy | <10cm                      | 6.9cm (hor.)<br>7.8cm (ver.) |
|               | #Points               | 43,971,092                 | 45,154,384                   |

Table 1. Detailed information of the test data

### 4.1 L-Junction Collecting and Matching

When matching the junctions on the aerial images of the two datasets,  $\mu_1$  (the threshold for junction-centre, see condition (2)) were set 3 pixels and  $\mu_2$  (the threshold for junction-branch, see condition (5)) were set 5 degrees. In the Guangzhou dataset, 42 L-junctions were manually collected and automatically matched. The screenshots of six L-junctions collected and matched in Guangzhou dataset are shown in Figure 8. The matching accuracies of the six illustrated junctions (Figure 8) are given in Table 2. In Ningbo dataset, 28 L-junctions were manually collected and automatically matched.



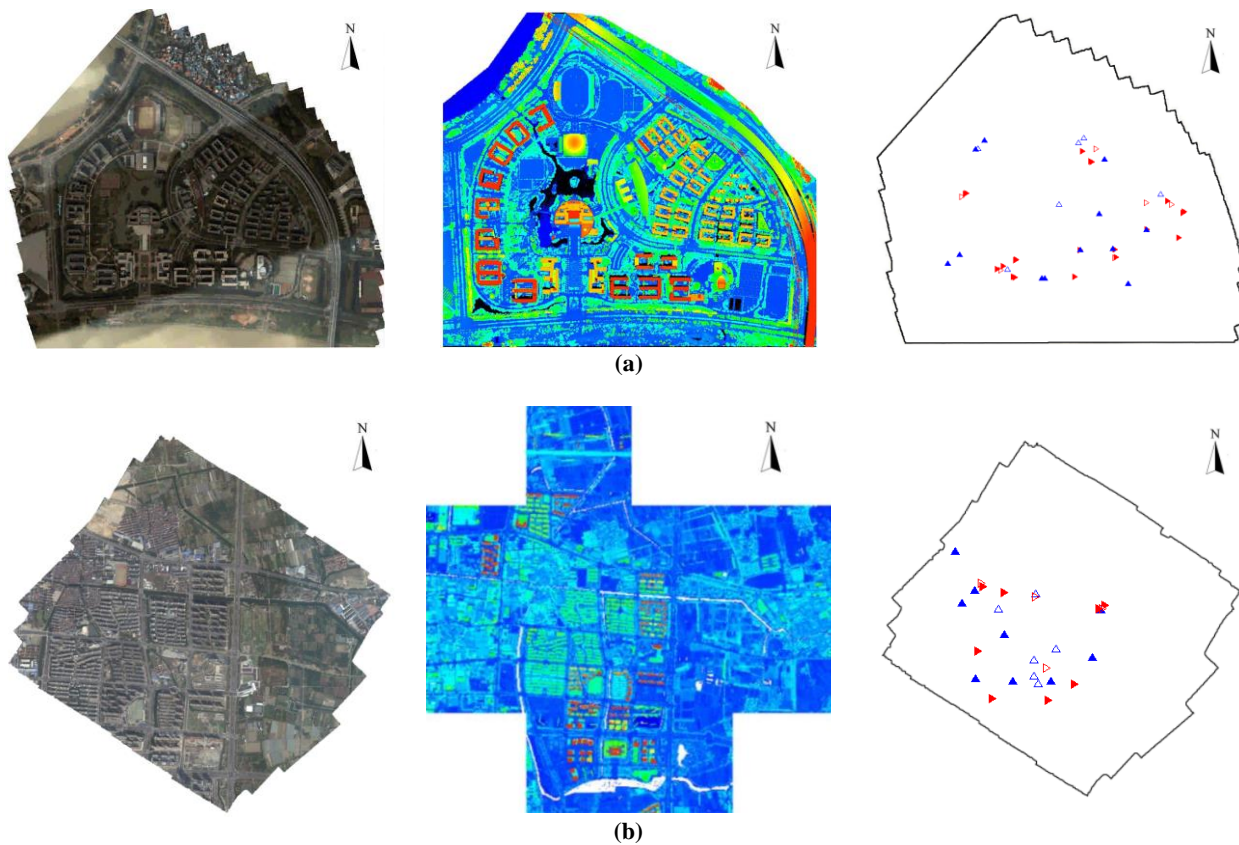


Figure 7. The ortho-maps, the LiDAR pointclouds, and the L-junction distributions (▲: HCJ, △: HICJ, ►: VCJ, ▷: VICJ) of (a) Guangzhou dataset and (b) Ningbo dataset.

| No. of Junction | RMS (pixel) of   |                  |                  | Maximum (pixel) of |                  |                  |
|-----------------|------------------|------------------|------------------|--------------------|------------------|------------------|
|                 | $\epsilon_{c,j}$ | $\epsilon_{p,j}$ | $\epsilon_{q,j}$ | $\epsilon_{c,j}$   | $\epsilon_{p,j}$ | $\epsilon_{q,j}$ |
| 1               | 0.86             | 0.43             | 0.55             | 1.61               | 0.95             | 1.10             |
| 2               | 1.09             | 0.90             | 0.64             | 1.63               | 2.02             | 1.35             |
| 3               | 0.96             | 0.62             | 0.40             | 1.91               | 1.62             | 1.14             |
| 4               | 0.89             | 0.66             | 0.51             | 1.69               | 1.30             | 1.36             |
| 5               | 0.88             | 0.92             | 0.32             | 1.95               | 2.03             | 0.87             |
| 6               | 0.67             | 0.58             | 0.27             | 1.66               | 1.28             | 0.58             |

Table 2. The matching accuracy of the junctions in Figure 8.

As is shown in Table 2, five out of six L-junctions had subpixel reprojection errors. All the L-junctions used in the two datasets had less-than-1.5-pixel RMS, among which 35 in Guangzhou, and 22 in Ningbo had subpixel RMS. Furthermore, the L-junction can even be accurately matched on the views where the junction-planes were invisible (see the images in row-1/column-5, row-2/column-2, row3/column-2, row-4/column-3, and row-4/column-5 in Figure 8). This is because the proposed matching strategy did not utilize the texture information of the images.

#### 4.2 3D Registration Accuracy

When searching the LiDAR points associated with an L-junction, the threshold  $\mu_H$  was set 3m. During the bundle block adjustment of the proposed approach, the threshold  $\mu_p$  was 0.15m. The method based on the closest-point principle and the collinearity equations (Huang, et al., 2018) was used for comparison. The tie-points used in the pre-process (POS-aided bundle adjustment) were also used in the 3D registration, so we did not need to obtain them again.

Figure 7 also illustrates the distributions of HCJs (Horizontal CJs), VCJs (Vertical CJs), HICJs (Horizontal ICJs), and VICJs (vertical ICJs) of both Guangzhou and Ningbo datasets. The number of each type of L-junctions is given in Table 3. The horizontal registration accuracy was evaluated with the VICJs (because they had vertical junction-planes with horizontal normal vectors) and similarly, the vertical registration accuracy was evaluated with the HICJs.

| Location  | #HCJs | #VCJs | #HICJs | #VICJs |
|-----------|-------|-------|--------|--------|
| Guangzhou | 12    | 18    | 6      | 6      |
| Ningbo    | 9     | 8     | 6      | 5      |

Table 3. The number of HCJs, VCJs, HICJs, and VICJs.

The registration accuracy is given in Table 4. As is shown, the L-junction based approach had better accuracy than Huang's method by the registration RMS of both the HICJs and the VICJs in both Guangzhou and Ningbo dataset. In Guangzhou dataset, Huang's method had lower maximum vertical accuracy (by HICJs) but higher maximum horizontal accuracy (by VICJs). In Ningbo dataset, the L-junction based approach achieved much better horizontal registration accuracy (by VICJs). The reason is that most of the LiDAR points in Ningbo dataset were on horizontal planes, thus the horizontal constraint was too weak for Huang's method. The L-junction based approach, on the other hand, can balance the constraints by controlling the numbers of HCJs and VCJs.

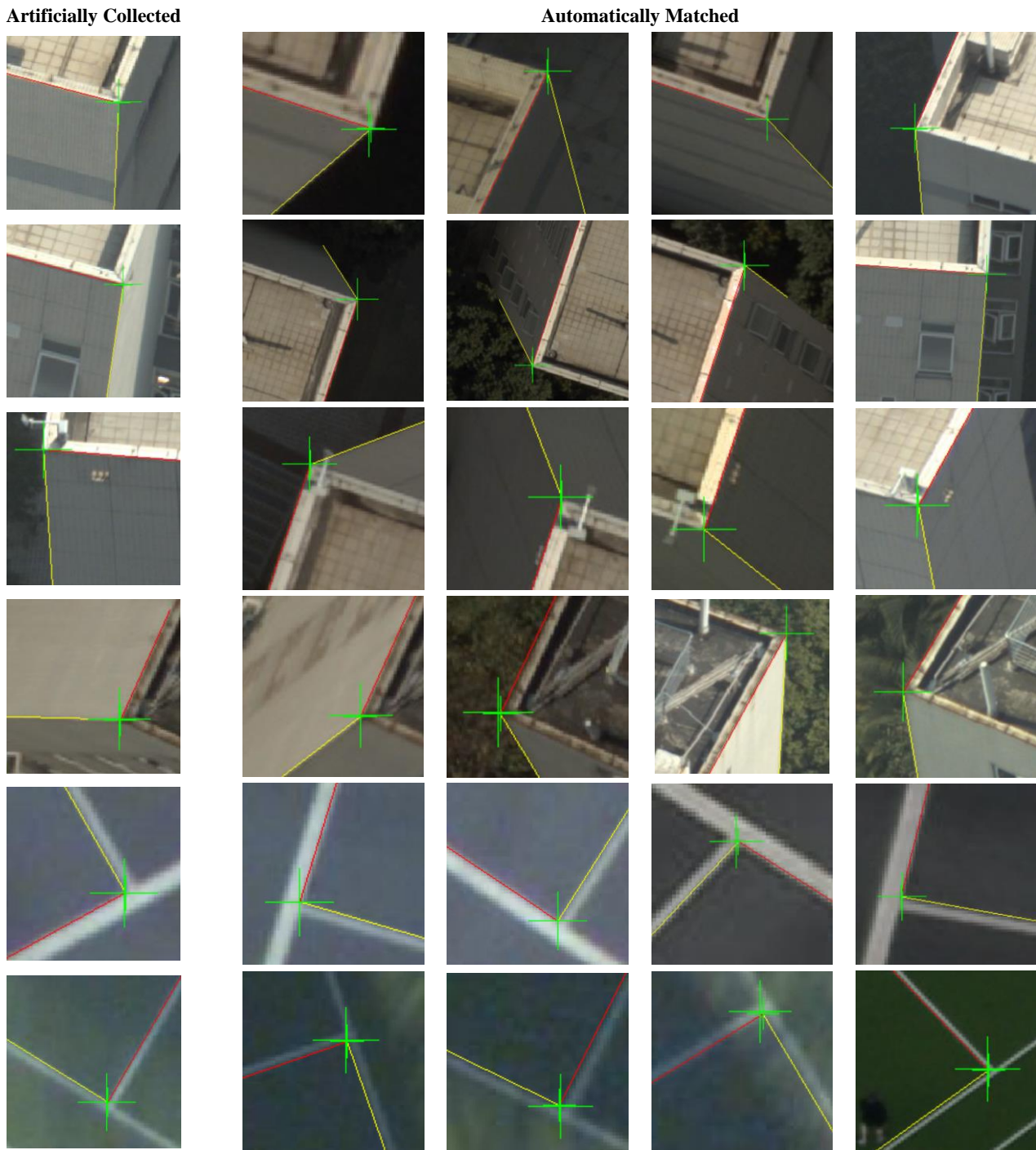


Figure 8. The screen-shots of the artificially collected L-junctions (the first column) and the automatically matched L-junctions.

| Registration Accuracy (cm) |     | Huang's Method |       | L-junction Method |             |
|----------------------------|-----|----------------|-------|-------------------|-------------|
|                            |     | VICJs          | HICJs | VICJs             | HICJs       |
| Guangzhou                  | RMS | 4.83           | 6.02  | <b>4.21</b>       | <b>5.72</b> |
|                            | max | <b>9.01</b>    | 10.44 | 11.53             | <b>9.92</b> |
| Ningbo                     | RMS | 8.80           | 5.58  | <b>4.46</b>       | <b>4.34</b> |
|                            | max | 17.90          | 14.81 | <b>9.90</b>       | <b>9.73</b> |

Table 4. The registration accuracy.

### 4.3 Discussion

With either Guangzhou dataset or Ningbo dataset, the whole process of the proposed approach took about half-an-hour manual works in searching the appropriate L-junctions for match, while other tasks (junction matching, LiDAR point searching, plane-constrained block adjustment, accuracy assessment) were fully automatic.

The junction matching accuracy was very high, the reprojection RMS of the junction-centre and the junction-branches can reach subpixel level for most L-junctions, which is very hard even for experienced workers. The experiment results of Guangzhou and

Ningbo datasets demonstrated that our L-junction based approach can reach very high accuracy in aligning the aerial images and LiDAR pointclouds.

## 5. CONCLUSION

In this paper, a novel L-junction based approach for semi-automatic accurate registration of the aerial images and the LiDAR pointclouds is proposed. The approach uses multi-view corresponding L-junctions to extract local planes and thus, form the planar constraint. The method of L-junction matching, using epipolar-geometry and vanishing-point constraints, achieves very high matching accuracy. The registration model is formed by adding the junction-projection costs and the plane-constraint costs to the conventional bundle block adjustment model. The registration accuracy in both horizontal and vertical directions is very high, being very close to the ranging accuracy of the LiDAR sensor and the image GSD. The proposed approach is demonstrated to be able to reach very high registration accuracy with very sparse plane-constraints (30 L-junctions for the 2415 images in Guangzhou and 17 L-junctions for the 1541 images in Ningbo), which means that users do not need to collect too many L-junctions manually.

The junction-based planar constraint is very suitable for aligning the very-high-resolution images to the sparse but accurate LiDAR data in urban areas, which is mostly the case of low-attitude (UAV/oblique) aerial photography and airborne LiDAR data. Because it is easy to find appropriate L-junctions in the urban area, and the junction-plane can be accurately extracted from the high-resolution images. For the rural areas, this approach is less suitable because selecting appropriate L-junctions will be very difficult. When aligning the aerial images with the terrestrial laser scanning data, the full-automatic iterative-closest-point (ICP) based method is a better choice because the point density will be very high and will not influence the accuracy.

The proposed approach treats the LiDAR points as reference data and adjust the image orientation parameters. However, with the junction-based planar constraints, one can also adjust the LiDAR orientation parameters when the aerial images are accurate enough to be used as reference data or adjust the orientation parameters for both the data sources concurrently when accurate ground control points are available. Besides, there are also many works to do to improve the proposed approach. Future works include:

1. Testing the different adjustment model with the junction-based planar constraints
2. Theoretically analysing the relationship between the junction-matching thresholds ( $\mu_1$  and  $\mu_2$ ) and the orientation accuracy of the aerial images;
3. Robust plane detection method that can eliminate the LiDAR points which do not belong to the junction-plane.
4. Automatically collecting the appropriate L-junctions lying on the real-existed flat object surfaces.

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under grant No. 41871368.

## REFERENCES

Akca, D., 2007. Matching of 3D surfaces and their intensities. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(2), 112-121.

Armenakis, C., Gao, Y. and Sohn, G., 2013. Co-registration of aerial photogrammetric and LiDAR point clouds in urban environments using automatic plane correspondence. *Applied Geomatics*, 5(2), 155-166.

Barsai, G., Yilmaz, A., Nagarajan, S. and Srestasathiern, P., 2017. Registration of Images To Lidar and GIS Data Without Establishing Explicit Correspondences. *Photogrammetric Engineering & Remote Sensing*, 83(10), 705-716.

Costa, F.A.L. and Mitishita, E.A., 2019. An approach to improve direct sensor orientation using the integration of photogrammetric and lidar datasets. *International Journal of Remote Sensing*, 40(14), 5651-5672.

Costa, F.A.L., Mitishita, E.A. and Centeno, J.A.S., 2017. A study of integration of LIDAR and photogrammetric data sets by indirect georeferencing and in situ camera calibration. *International Journal of Image and Data Fusion*, 8(2), 94-111.

Gruen, A. and Akca, D., 2005. Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3), 151-174.

Huang J, Zhang X, Xin Q, et al. 2019. Automatic building extraction from high-resolution aerial images and LiDAR data using gated residual refinement network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 151(1), 91-105.

Huang, R., Zheng, S. and Hu, K., 2018. Registration of Aerial Optical Images with LiDAR Data Using the Closest Point Principle and Collinearity Equations. *Sensors*, 18(6), 1770.

Mishra R, Zhang Y. 2012, A review of optical imagery and airborne LiDAR data registration methods. *The Open Remote Sensing Journal*, 5(1), 54-63.

Peng S, Zhang L. 2019. Automatic registration of optical images with airborne LiDAR point cloud in urban scenes based on line-point similarity invariant and extended collinearity equations. *Sensors*, 19(5), 1086-1103.

Shorter N, Kasparis T. 2008. Autonomous registration of lidar data to single aerial image. *IGARSS 2008. IEEE Int 2008*, 5, 216-19.

Wong, A. and Orchard, J., 2008. Efficient FFT-accelerated approach to invariant optical-LIDAR registration. *IEEE transactions on geoscience and remote sensing*, 46(11), 3917-3925.

Zhang J, Lin X. 2016. Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *International Journal of Image and Data Fusion*, 8(1), 1-31.

Zheng, S., Huang, R. and Zhou, Y., 2013. Registration of optical images with LiDAR data and its accuracy assessment. *Photogrammetric Engineering & Remote Sensing*, 79(8), 731-741.