

VERTICAL ORIENTATION CORRECTION OF UAV IMAGE-BASED POINT CLOUDS USING STATISTICAL MODELING OF GABLE ROOF GEOMETRY

P. Polewski¹, W. Yao^{1,*}, L. Fang¹

¹ Dept. of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong -
{przemyslaw.polewski,wei.hn.yao,li.fang}@polyu.edu.hk

KEY WORDS: 3D shape fitting, parameter voting, coregistration, gable roofs, rural area

ABSTRACT:

Coregistration of point clouds obtained from various sensors is an important part of workflows for automatic building reconstruction from remote sensing data. Many approaches assume a common Z axis between the coordinate systems, and perform coregistration in 2D. While this assumption is usually valid for laser scanning (LS) data, for photogrammetric point clouds the Z axis is in general different from the world Z axis, and requires correction e.g. by manually measured ground control points (GCP). In this paper, we propose a fully automatic, GCP-free procedure for finding the world Z axis in rural areas, based on the relationships of planar surfaces in building gable roofs. Instead of performing direct gable line detection, we derive these lines as theoretical intersections between adjacent roof planes from 3D shape fitting. Each gable roof then casts a vote for both the Z axis direction and sign based on roof convexity constraints, and the votes are aggregated through a non-parametric kernel density estimator model. Experiments on two real world UAV image-based point clouds show that the Z axis recovered by our method leads to high-accuracy planimetric coregistration, with a median distance over 89 as well as 149 matched linear feature pairs (respectively for dataset 1 and 2) lying below 1 cm. Our results indicate that a high-quality vertical orientation can be achieved without using any GNSS or IMU hardware, which enables the use of low-cost UAV platforms for suburban and rural mapping tasks.

1. INTRODUCTION

Automatic mapping of urban and rural areas from remote sensing data is a topic of significant practical importance as a means for obtaining building models and related cadastral maps (Jazayeri et al., 2014, Sun, Salvaggio, 2013). In recent years, UAV-based camera systems are increasing in popularity as an alternative to laser scanning for urban mapping (Nex, Remondino, 2014). The main advantage of such UAV-mounted systems is their significantly lower cost and hence accessibility to a wider circle of institutions and individuals. However, due to the lack of prohibitively expensive odometry hardware, photogrammetrically reconstructed point clouds originating in such systems are scaled and oriented arbitrarily. This is in contrast to laser scanning data, which is either already georeferenced (as in aerial laser scanning), or at least properly scaled and aligned with the world Z coordinate axis (for terrestrial laser scanning). In order to perform fundamental cadastral mapping tasks such as building outline extraction and planimetric area calculation, the Z axis must be properly aligned and the point coordinates should be at scale. The alignment of the vertical axes is also a requirement for many point cloud coregistration methods (e.g. (Von Hansen et al., 2008, Yang et al., 2015)), which must be carried out if change detection or comparison to previously acquired data is to be conducted.

One way of properly orienting and scaling such photogrammetric data is through the use of manually measured ground control points (GCPs), however such a procedure is time consuming and expensive, which defeats the purpose of applying a cheaper platform. In this paper, we propose an alternative method of automatically finding the world Z axis orientation for photogrammetric points containing

buildings with gable roofs, in case no ground control points are available. Specifically, we find matching pairs of roof planes and utilize relationships between their surface normals as well as convexity constraints to derive a vote for the global Z axis for each gable roof. The votes are aggregated in a kernel density estimator model, whose global maximum represents the optimal Z axis (Polewski et al., 2017). Our method is insensitive to point cloud scale, as all parameters concern either relative distances or angular deviations. Once the Z axis is corrected, the scale may be determined by coregistering the point cloud to another dataset with known scale, or by comparing building dimensions measured on the (now properly oriented) orthophotograph with previous cadastral records or other measurements. We show that the analytically computed gable roof intersection lines provide a natural feature source for high-accuracy coregistration.

2. RELATED WORK

There exists little prior work on the automatic orientation of photogrammetric point clouds. In (Gerke, 2011), additional structural information is integrated into the bundle adjustment procedure to reduce the number of GCPs necessary for orienting imagery of urban scenes. However, a number of GCPs is still required to complete this task. One of the relevant works might be found in (Roh et al., 2017), which developed an image-based correction method for heading orientation in mapping and navigation applications for urban canyon. The approach utilized the direction of the structural edges from buildings to restrict the estimation error. However, a number of important works concerning relevant topic of co-registration for point clouds have been proposed before. In (Von Hansen et al., 2008), the vertical axes are assumed to be aligned, and the target transformation parameters are the scale, 3D translation and rotation angle around the Z axis. The rotation

*Corresponding author

is estimated from correlating histograms of the line orientations in both point clouds. This is based on a strong assumption of rectangular buildings located along parallel streets, resulting in distinctive local maxima in the histograms. While this may be valid in strongly built-up urban locations, in rural areas the streets do not always follow this pattern (see Fig. 3(a)). The method of Wu et al. (Wu et al., 2014) also assumes vertical axis alignment. The authors find the transform which minimizes the deviations between the reference and fitted linear features of corresponding buildings. Novak and Schindler (Novak, Schindler, 2013) propose a method for coregistration where the point clouds' Z axes are allowed to diverge. The principal point cloud axis is then approximated by dominant surface normal orientations. While this approach may yield useful axis estimates for point clouds representing isolated objects or urban regions with buildings featuring flat roofs, we argue that in case of rural scenes dominated by gable roofs and acquired from a UAV perspective, it may be insufficient. This is because few surface normals in such a scene would be oriented according to the world Z axis, especially if the terrain is covered with vegetation or is not flat. To deal with this issue, we propose a Z axis voting scheme based on the intersection lines of adjacent gable roof planes.

3. PLANAR SURFACE SEGMENTATION

Our strategy for recovering the point cloud's Z axis is based on intersection lines of adjacent planes. We do not attempt to detect such lines directly within the 3D data, since they might not be accurately represented due to the presence of outliers and/or noise. Instead, we pursue a more robust approach of precisely determining the plane parameters of neighboring components and analytically calculating their intersection lines. To ensure that the plane detection method is not only translation- and rotation- but also scale-invariant, we designed the processing pipeline to use only parameters relating to relative distances, angles, and nearest neighbor counts, avoiding fixed absolute distances.

3.1 Data smoothing

Photogrammetrically generated point clouds may contain more noise compared to their counterparts obtained by laser scanning due to artifacts arising from the dense matching algorithm. To alleviate this problem, we perform a moving least squares-based correction of the point cloud (Levin, 2004). In the general formulation, each point p is replaced by its projection $p' = Q'(p)$ onto the m -th order polygonal approximation Q of the local surface around p having the minimal least-squares residual:

$$Q' = \operatorname{argmin}_Q \sum_{i \in I} \|Q(p_i) - p_i\|_2^2 \theta_i \quad (1)$$

In the above, I represents some index set defining p 's neighborhood, whereas θ represents a non-negative weight function. Since the target surface type is planar, it is natural to use a first-order approximation. Also, although the index set I is commonly defined through a spherical neighborhood of radius d_n centered around p , this would violate our assumed policy of avoiding absolute distance parameters. Therefore, we instead apply a k -nearest neighborhood rule $N_k(p)$. Assuming a Hessian normal form for the plane $Q : n \cdot x - D = 0$, where n, D represent the respectively plane's surface normal and distance

to origin, we can rewrite the above equation more specifically as:

$$Q' = \operatorname{argmin}_Q \sum_{i \in N_k(p)} (n \cdot p_i - D)^2 \theta_i \quad (2)$$

This procedure is equivalent to projecting p onto its best-fit (least-squares) local plane defined by p 's k nearest neighbors.

3.2 Plane fitting

We apply the region growing algorithm (Rabbani et al., 2006) to the point cloud in order to partition it into spatially contiguous planar components. After the surface smoothing, the point normals are more coherent, which allows us to apply a strict smoothness constraint α_{thr} to ensure that each connected component contains only one planar surface. Specifically, α_{thr} represents the maximum permitted angular deviation between a candidate point's surface normal and the current normal estimate of the whole planar component. If this angle does not exceed α_{thr} , the candidate point is accepted and the component's normal estimate is updated, else the point is discarded. Similar to the case of surface smoothing, we utilize a nearest- k neighborhood structure instead of a predefined neighborhood radius. Each segmented planar component with at least N_{thr} member points undergoes least-squares plane fitting, recovering the plane normal and distance from origin.

3.3 Rectangular outline calculation

In order to determine whether two planes intersect, our method requires some notion of plane proximity. Since the scale is unknown, we cannot set a maximum absolute distance threshold between member points of the considered planar clusters. Instead, we define plane proximity based on spatial relationships between their rectangular outlines. Specifically, we project all the planar cluster's points onto the corresponding plane obtained from least-squares fitting, and perform principal component analysis of the projected coordinates. By retaining only the first 2 principal components with largest variance, we obtain a new two-dimensional basis B' spanning the current plane. We then compute the minimum-area 2D oriented bounding box (Toussaint, 1983) of the plane's member points, expressed in their planar coordinates (in B'). Finally, the 4 vertices of the bounding box are transformed from B' into the original 3D coordinate system.

4. INTERSECTION LINE-BASED AXIS VOTING

4.1 Deriving Z axis votes from gable roofs

To find gable roof intersection lines, we examine all pairs of planar components and check if the angle between the normals of their fitted planes lies in the range of valid gable roof angles, $[\alpha_{min}; \alpha_{max}]$. It is assumed that $\alpha_{max} < 85^\circ$ to avoid confusion with walls. If two planes P_i, P_j meet this condition, the proximity of their bounding boxes is examined. Each pair of bounding box edges $e_k^i, e_l^j, 1 \leq k, l \leq 4$ is checked against 3 criteria. Letting $\prod_k^i(e)$ and $|e|$ denote respectively the projection of edge e onto edge e_k^i , and the length of e , the criteria can be expressed as:

- Intersection angle between e_k^i, e_l^j should not exceed α_{thr}^g
- The ratio of edge length covered by projecting the edges onto each other, i.e. $\max(|\prod_k^i(e_l^j)|/|e_l^j|, |\prod_l^j(e_k^i)|/|e_k^i|)$, should be at least r_{prj}

- The average projected distance between equally-spaced intervals on e_k^i, e_l^j should not exceed $d_{thr}^g \min(|e_k^i|, |e_l^j|)$ (see Fig. 1)

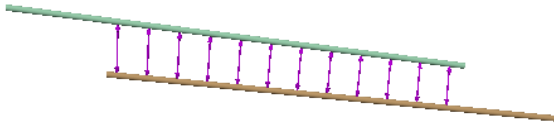


Figure 1. Distance between equidistant points on the first line, and their projections on the second line.

Plane pairs P_i, P_j with a pair of bounding box edges fulfilling all 3 criteria are considered as forming a gable roof. Next, a gable intersection line is computed analytically based on the parameters of the intersecting planes. Let $(n^i, D^i), (n^j, D^j)$ denote the respective fitted plane parameters. The direction vector of the intersection line can be recovered as the cross-product of the normals: $v^{ij} = n^i \times n^j$. To find a point p^{ij} on the line, we determine the intersection point of both P_i, P_j and one of the planes XY, XZ, YZ , spanned by pairs of the coordinate axes. This can be achieved by setting the appropriate coordinate to 0, and solving the obtained linear system with 2 equations and 2 variables. For example, to find the intersection with the YZ plane, we set $p_x^{ij} = 0$ and solve:

$$\begin{pmatrix} n_y^i & n_z^i \\ n_y^j & n_z^j \end{pmatrix} \begin{pmatrix} p_y^{ij} \\ p_z^{ij} \end{pmatrix} = \begin{pmatrix} -D^i \\ -D^j \end{pmatrix} \quad (3)$$

The procedure for the remaining two planes can be defined in a similar manner. The choice of the specific plane depends on the line orientation v^{ij} .

In the next stage, we wish to extract the gable roof's vote for the Z axis. If the orientations of all the plane normals were guaranteed to be consistent, this Z axis estimate would simply be the average of the two (unit-length) normals: $0.5(n^i + n^j)$. However, there is no unambiguous way to define the normal orientation (sign) during plane fitting, since the point cloud is rotated arbitrarily. As Fig. 2(b) shows, this may lead to averaging two normal vectors having opposite orientations with respect to the direction of the gable roof's convexity. In particular, when one of n^i, n^j is facing downwards of the gable roof, and the other one upwards, then their average yields a vector Z_2 which is not a correct estimate of the Z axis. To ensure the proper relative normal orientations, we perform the following procedure. First, a planar coordinate system is defined, associated with the plane π having v^{ij} as its normal vector (see Fig. 2(a)), and centered at the projection O of p^{ij} onto π . Let $c^{i'}, c^{j'}$ denote the centroid projections (onto π) of the bounding boxes belonging to P_i, P_j . We examine whether traversing the point sequence $c^{i'}, O, c^{j'}$ constitutes a left turn or a right turn by checking the sign of the cross product $(O - c^{i'}) \times (c^{j'} - O)$. If the normals of P_i, P_j are consistent, traversing n^i followed by n^j should produce a turn in the same direction. If this is not the case, it means that n^i and n^j lie on opposite sides of the gable roof. In this case, we flip the sign of n^i . After this correction, averaging $Z_e = 0.5(n^i + n^j)$ is guaranteed to yield the correct Z axis estimate (Z_1 from Fig. 2(b)), up to the sign of the vector. To make sure that we obtain a Z axis pointing 'upwards', we consider the angle between Z_e and $(O - c^{i'})$, and pick the sign of Z_e which minimizes this angle.

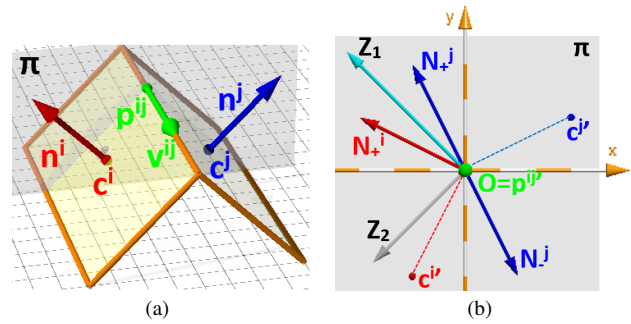


Figure 2. (a) Pair of planar facets i, j forming a gable roof. The direction vector v^{ij} of the intersection line is perpendicular plane π 's normal. Normal vectors and bounding box centroids of roof parts denoted by n^i, n^j and c^i, c^j . (b) Normal vectors with two orientations (+, -) and centroids projected onto π . Only averaging two normal vectors with 'positive' orientation (according to the roof convexity $c^{i'}Oc^{j'}$) will result in the proper Z axis estimate Z_1 .

4.2 Vote aggregation

Each detected plane pair forming a gable roof casts a single vote on both the direction (line) and orientation (sign) of the whole dataset's Z axis. Since we do not perform any filtering or classification of the planes based on spatial or visual characteristics, it is possible that false positive pairs of invalid objects detected as gable roofs may cast their (wrong) votes. To mitigate this, we apply a vote aggregation strategy which is insensitive to outliers and instead will extract the Z axis associated with the mode of the vote distribution. Specifically, we use the method described in (Polewski et al., 2017), first projecting the 3D axis votes (x, y, z) onto the Gaussian sphere and thus obtaining a 2D representation (with unit radius) in terms of angles:

$$\begin{cases} \theta = \tan^{-1} \frac{y}{x} \\ \phi = \cos^{-1} z \end{cases} \quad (4)$$

To obtain the most probable Z axis line, a kernel density estimator model is constructed in the 2D vote space, and the global maximum is found using multiple-restart gradient ascent search. Finally, the sign is determined by majority voting among the partial Z axis estimates (votes) which lie near the KDE's global optimum.

5. EXPERIMENTS

5.1 Material

To assess the applicability of the proposed method, two UAV flight campaigns were performed in a rural area close to Wuhan, China, containing mostly residential buildings with gable roofs. A dataset of 130 images of 5456x3632 pixels was acquired by a Sony ILCA QX1 camera, with a GSD of 2cm and a view angle of 45°. The camera was mounted on a Tianxing Octocopter HO1300+ from South Surveying Ltd. with a total payload of 7 kg. The flight was planned for the images to have a forward and a side overlap of 80% and 60% respectively owing to a flight height of 90 m. The platform included precise GNSS and IMU hardware, yielding accurate trajectory data for georeferencing.



Figure 3. Orthophotographs depicting parts of (a) Plot 1 and (b) Plot 2. Green and red lines indicate, respectively, successfully matched and unmatched gable lines between the georeferenced and transformed datasets. Cyan and orange rectangles denote matched and unmatched roof plane facets.

State-of-the-art software was used for structure-from-motion based photogrammetric generation of point clouds from the imagery. For each of the two plots, two sets of point clouds were generated: (i) properly scaled and oriented, based on the recorded sensor trajectory, and (ii) unreferenced, generated using only tiepoints derived from the imagery, resulting in a random scale and orientation.

5.2 Experimental setup and evaluation

Our method was applied to the unreferenced version of each plot's point cloud, and the cloud was rotated according to the obtained corrected Z axis. As the 'true' Z axis is unknown, we instead focused on comparing planimetric features between the georeferenced and automatically corrected point clouds after projection on the XY plane. DTM filtering was applied to remove ground points. We then calculated the gable roof intersection lines for both datasets. For each plot, we manually chose 20 corresponding pairs of buildings, and utilized the midpoints of their gable lines as tiepoints to calculate the optimal 2D transformation consisting of a scaling, rotation and offset, from the unreferenced coordinate system to the referenced one, using the method of Umeyama (Umeyama, 1991). After coordinate transformation, we matched all pairs of gable lines having an angular deviation below 1° and average distance below 5 cm. Matching quality was assessed based on the mean angular and spatial deviations between the matched lines. Moreover, we computed approximate roof areas as rectangular bounding boxes aligned with the respective gable intersection line. The roof areas of corresponding buildings were compared between the original (referenced) and corrected/transformed point clouds for both plots. Two roof planes were considered as matched if their 2D intersection area constituted at least 0.7 of each individual area, and the deviation in their (projected) areas was below 15%, with respect to the smaller polygon.

6. RESULTS AND DISCUSSION

Table 1 shows the accuracy of the gable line and roof plane planimetric matching after the Z axis correction is applied. For both plots, the spatial and angular distances between corresponding lines are low, with medians lying below 1 cm and 0.1° , which represents sub-pixel accuracy with respect to the camera GSD of 2 cm. Despite quite strict matching criteria, 89 and 147 gable lines could be matched in Plot 1

Table 1. Accuracy of gable line and roof polygon matching. μ, σ, m denote respectively the mean, standard deviation and median of the error metric.

	Plot 1			Plot 2		
	μ	σ	m	μ	σ	m
Gable lines						
Angular deviation [$^\circ$]	0.13	0.18	0.07	0.14	0.22	0.06
Line distance [mm]	11.6	9.5	9.0	10.9	9.5	7.6
Roof areas						
Relative error [%]	3.6	3.2	2.6	4.0	3.1	3.4
Absolute error [m ²]	1.3	1.3	1.0	1.4	1.4	1.0

and 2 (see Fig. 3), which suggests that the gable line positions can be calculated with high accuracy using the analytical formulation. We believe this is associated with the fact that the plane parameters can be determined more robustly based on a larger quantity of its inliers, compared to explicit fitting of the intersection line. To estimate the precision of the axis estimation, we simulated a tilt of the corrected Z axis by 1 to 5° in a random direction, recalculating the Umeyama transformation and re-evaluating the matched line distance. For each angle, the random rotation was repeated 10 times, and the least-distance result was recorded (see Fig. 4(a)). Not a single of the 50 conducted rotations led to an improvement of the distances in either Plot 1 or Plot 2. Tilting the axis by even 1° resulted in more than doubling the average line distance, suggesting that the Z axis deviation between the georeferenced and corrected datasets is below 1° . The high-accuracy co-registration based on gable lines could be exploited for finding the scale of the unreferenced point cloud, by co-registering with another point cloud or vector model with known scale. In our case, the scales recovered by the Umeyama transformation were 11.96 and 17.05 respectively for Plot 1, 2.

Figure 4(b) presents the distribution of the relative area deviations between the matched planar roof facets. In Plot 1, a total of 207 and 208 planes were detected respectively in the georeferenced and raw point clouds, out of which 162 were considered as matched pairs. It is worth noting that despite the scale factor of nearly 12 and the fact that these two point clouds were generated separately under different conditions (camera reference positions), more than 42% of the matched planes had a relative error below 2%, and more than 23% below 1%,

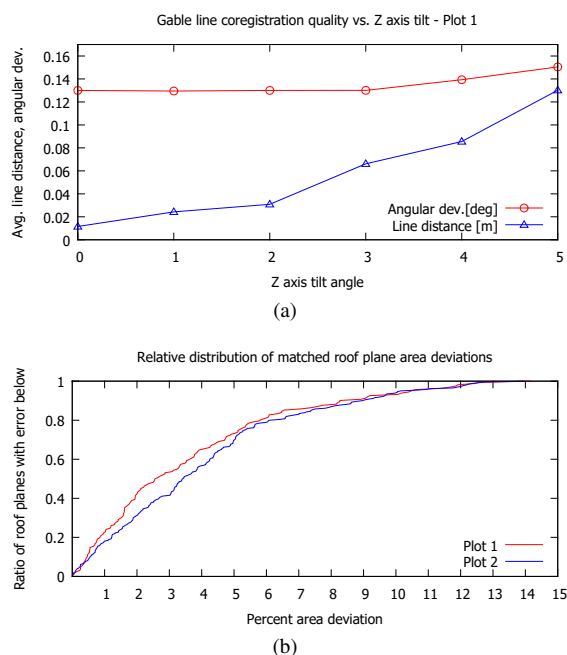


Figure 4. (a) Mean distances and angular deviations between corresponding gable lines as a function of Z axis tilt (in degrees) - Plot 1. (b) Distribution of relative roof area deviations between referenced and unreferenced datasets.

resulting in a median error of 2.6%. Plot 2 turned out to be slightly more complex, with 260 out of 384 (georeferenced) and 355 (raw) planes matched, yielding a median error of 3.4%. This time, the percentage of roof planes with a relative error below 2% was 32%, which is probably associated with the higher number of smaller roof parts belonging to utility structures/sheds etc. We believe that the unmatched portions of the roof facets may be explained by the differences in the point clouds themselves (due to presence/lack of georeferencing) which propagate to further processing steps (smoothing, region growing, plane fitting). Also, the set of buildings in the pair of processed point clouds was not identical due to differences in DTM height filtering between the georeferenced and raw data, since the filtering was applied before proper scaling of the latter, utilizing an approximate filtering height different from the scaled setting.

7. CONCLUSIONS

This paper presented an automatic method for determining the vertical orientation of ungeoreferenced, photogrammetrically generated point clouds depicting rural area scenes. The method utilizes statistical modeling of gable roof geometry to recover the most probable Z axis. Experiments showed that the resulting vertical orientation has a precision below 1 degree, compared to a georeferenced dataset covering the same scene. Moreover, once the Z axis is corrected, the gable roof intersection lines were found to be a good source of features for planimetric coregistration, achieving a relative matched linear feature distance below 1 cm, or half the data acquisition sensor's GSD. Our results show that a high-accuracy vertical orientation of a UAV-based point cloud is possible without using GNSS/IMU hardware. Also, the methods could be extended towards a fully automatic co-registration between

both point clouds and vector 3D models of rural areas, provided that a scheme for finding corresponding linear features is defined.

ACKNOWLEDGEMENTS

The work described in this paper was supported by a grant from the Germany / Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the German Academic Exchange Service (Reference No.G-PolyU502/17).

REFERENCES

- Gerke, Markus, 2011. Using horizontal and vertical building structure to constrain indirect sensor orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 307 - 316.
- Jazayeri, Ida, Rajabifard, Abbas, Kalantari, Mohsen, 2014. A geometric and semantic evaluation of 3D data sourcing methods for land and property information. *Land Use Policy*, 36, 219 - 230.
- Levin, David, 2004. Mesh-independent surface interpolation. G. Brunnert, B. Hamann, H. Müller, L. Linsen (eds), *Geometric Modeling for Scientific Visualization*, Springer Berlin Heidelberg, Berlin, Heidelberg, 37–49.
- Nex, Francesco, Remondino, Fabio, 2014. UAV for 3D mapping applications: a review. *Applied Geomatics*, 6, 1–15.
- Novak, D., Schindler, K., 2013. Approximate registration of point clouds with large scale differences. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W2, 211–216.
- Polewski, Przemyslaw, Yao, Wei, Heurich, Marco, Krzystek, Peter, Stilla, Uwe, 2017. A voting-based statistical cylinder detection framework applied to fallen tree mapping in terrestrial laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129, 118 - 130.
- Rabbani, T., van den Heuvel, F. A., Vosselmann, G., 2006. Segmentation of point clouds using smoothness constraint. *IEVM06*.
- Roh, H., Jeong, J., Kim, A., 2017. Aerial Image Based Heading Correction for Large Scale SLAM in an Urban Canyon. *IEEE Robotics and Automation Letters*, 2, 2232-2239.
- Sun, S., Salvaggio, C., 2013. Aerial 3D Building Detection and Modeling From Airborne LiDAR Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6, 1440-1449.
- Toussaint, Godfried, 1983. Solving Geometric Problems with the Rotating Calipers. In *Proc. IEEE MELECON '83*, 10—02.
- Umeyama, S., 1991. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, 376-380.
- Von Hansen, Wolfgang, Gross, Hermann, Thoennessen, Ulrich, 2008. Line-based registration of terrestrial and airborne LIDAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 37, 161–166.
- Wu, Hangbin, Scaioni, Marco, Li, Hanyan, Li, Nan, Lu, Minfeng, Liu, Chun, 2014. Feature-constrained registration of building point clouds acquired by terrestrial and airborne laser scanners. *Journal of Applied Remote Sensing*, 8, 8 - 8 - 22.
- Yang, Bisheng, Zang, Yufu, Dong, Zhen, Huang, Ronggang, 2015. An automated method to register airborne and terrestrial laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 109, 62 - 76.