

Accurate Geo-Localization of Low-Altitude Overhead Images From 3D Point Clouds

Yuhang Li

University of Science and Technology of China
P.O. Box 4, Hefei, Anhui, 230026, China
lyh9001@mail.ustc.edu.cn

Xuejin Chen

University of Science and Technology of China
P.O. Box 4, Hefei, Anhui, 230026, China
xjchen99@ustc.edu.cn



Figure 1: (a) The overhead image is captured by an aerial device in a low altitude. (b) The point cloud is scanned by a laser device on the ground. We extract the roof contours (in different colors) according to the altitude histogram of points (c) The contours are matched with the overhead image respectively to achieve an initialization for optimizing the global matrix. (d) The camera pose is estimated after the iterative global matrix optimization. We project the contours on the overhead image to show the results.

ABSTRACT

Geo-localization, aiming at register images with 3D models, is a key technique to many applications, such as image-based navigation, augmented reality, 3D city modeling, etc. We present a geo-localization method by estimating the camera pose of a low-altitude overhead image with respect to a corresponding building point cloud. The observations that vertical facades of a point cloud correspond to edges of building roofs in the overhead image and that roofs at different altitudes are of different scales in the overhead image inspire us to treat this geo-localization problem as a combination of a multi-layer shape matching problem and a global optimizing procedure. We test the proposed approach on a newly collected dataset. (Xuejin: What newly collected? no existing dataset? what kind of dataset? Describe more..)

CCS CONCEPTS

- Computing methodologies → Camera calibration; Shape inference; Point-based models;

KEYWORDS

Geo-localization, point cloud, shape matching

ACM Reference format:

Yuhang Li and Xuejin Chen. 2017. Accurate Geo-Localization of Low-Altitude Overhead Images From 3D Point Clouds. In *Proceedings of SIGGRAPH 2017 Poster, Los Angeles, CA, USA, August 2017*, 3 pages.

DOI: 10.475/123_4

1 INTRODUCTION

As a key technique to image-based navigation, augmented reality, 3D city modeling and so on, geo-localization has drawn massive attentions in the literature. The geo-localization problem is defined as aligning an image with a 3D model to estimate the camera pose with respect to the global coordinate of the 3D model. (Xuejin: of what? Whose global coordinate?). The sensor prior, including the compass and GPS, (Xuejin: including what? be specific) is optionally employed.

(Xuejin: Then talking about previous methods.) There are two typical types of images are used in geo-localization systems. One is the facade-view image captured on the ground [Arth et al. 2015]. The other one is remote sensing image captured in significantly high altitude [Liu et al. 2010; Zhang et al. 2014]. (Xuejin: Do we need more references?) (Xuejin: Discuss more about their methods.)

(Xuejin: Our method) (Xuejin: What is the motivation of using overhead-image?) With the populization of consumer-level drones, low-altitude overhead images can be easily obtained for buildings. In this work, we estimate the camera pose using an overhead image captured by a low-altitude aerial device as query and a corresponding building point cloud as 3D model. The building point clouds we deal with are scanned by a laser device, and only facades of buildings are able to be scanned. Comparing to existing methods using images captured on the ground [Arth et al. 2015] or in high altitude [Ni et al. 2013; Zhang et al. 2014], we face challenges that we are not able to take advantages of vanishing points in overhead image and suffer from more critical perspective effect in low-altitude image.

To build the correspondence between the overhead image and the point cloud scanned on the ground, we observe that: 1) vertical facades of a point cloud correspond to edges of building roofs in the overhead image; 2) roofs of different altitudes are in different scales in the overhead image due to perspective projection. Based

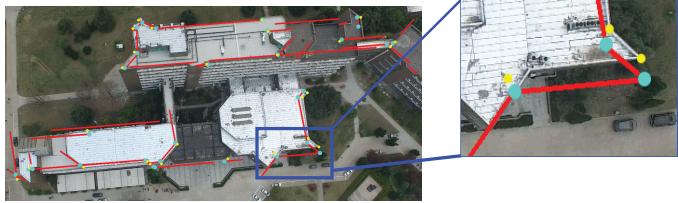


Figure 2: Finding paired 3D and 2D feature points: for a 3D feature point, we project it on the overhead image (magenta points) using current project matrix and search in its $k \times k$ neighborhood for a corner (cyan points). These 3D feature points and corresponding corners form pairs of feature points for next iteration of calculating project matrix.

on these two observations, we solve this geo-localization problem with a combination of a multi-layer shape matching problem and a global optimization of camera pose estimation.

(Xuejin: Our method is able to handle various types of scenes, including .. overhead images containing partial buildings.)

2 OUR APPROACH

Given an overhead image and the point cloud of a building, our method consists of two parts. The first part is a shape matching between the 3D contours of the point cloud at multiple sampled altitudes and the edges in the overhead image, as Figure 1(b)(c) show. The second part is a global optimization of the camera pose by refining point correspondences based on the shape matching results, as Figure 1(d) shows.

Shape Matching. The vertical directions of point clouds we use are aligned to the gravity direction in the reality. If this is not satisfied, a technique of finding the up-vector based on principle component analysis (PCA) [Ni et al. 2013] can be used to rotate the point cloud to meet the alignment. We divide the point cloud into several part vertically by sampled altitudes determined by sharp changes in the altitude histogram of points. (Xuejin: The sampling interval is determined ... as what?) In each interval, points are vertically projected to the upper cross section and fitted into 2D line segments on the cross section using a RANSAC line fitting algorithm. (Xuejin: using the ... algorithm). Taking line segments at each altitude as a contour of a building roof, we match contours with the edge map of the overhead image respectively using a shape matching technique [Zhang et al. 2014]. (Xuejin: Directly use the shape method without any modification? The same matching objective function with the reference?) A local project matrix is achieved for each contour after shape matching. The project matrix is defined as $M_0 = K_0[R_0|T_0]$, where K_0 is the pre-calibrated intrinsic camera matrix, R_0 is a 3×3 rotation matrix, and T is a 3-D translation vector. Here R_0 is a 1 degree-of-freedom (dof) matrix since we assume that the optical axis of the camera is aligned to the vertical direction of the point cloud and only the rotation of the vertical direction is needed to be estimated. The assumption will be discarded in global optimization part.

From Figure 1(c), we can see that the 2D shape matching generates a rough alignment between the contours and the overhead

image. However, there are obvious deviation at corners due to the non-ideal perpendicular projection of the low-altitude camera. Therefore, we need a global project matrix instead of local ones to estimate the camera pose. So we treat the results of shape matching as the initialization of following global optimizing procedure.

Global Optimization. We define the global project matrix from the 3D point cloud to the image as $M = K[R|T]$, where K is the intrinsic camera matrix, R is a 3 dof rotation matrix, and T is a 3-D translation vector. To calculate the global project matrix, we first find a set of paired 3D feature points and corresponding 2D feature points. For 3D feature points, we take the intersections of adjacent line segments of the contours. And then we employ an iterative algorithm to optimize the global project matrix by alternatively finding corresponding 2D feature points and calculating the project matrix, where we utilize the results of shape matching as initial project matrix.

To be more specific, we apply two steps in every iteration alternatively until the convergence. 1) As shown in Figure 2, for a 3D feature point x we project it on the overhead image using current project matrix to get a 2D point q (Cyan points). We then search its $k \times k$ neighborhood for a 2D corner p with the maximum Harris corner response (yellow points). We set $k = 31$ in all our experiments. 2) With a set of paired 3D and 2D feature points, we can optimize the global project matrix M by minimizing the average of distances of pixels between projected contours and the edges of the overhead image, which can be accelerated by distance transformation. And the newly calculated project matrix is used for the next iteration of finding paired feature points.

With the optimized global project matrix and focal length of the camera, we can finally estimate the 6 dof camera pose of the overhead image in point cloud coordinate system. (Xuejin: So you do not refine the camera intrinsic parameters?) We need focal length to determine the translation of z .

3 EXPERIMENTS AND FUTURE WORK

We tested the proposed approach on wide range of buildings. We scanned 10 buildings in our campus using a laser 3D scanner. A group of overhead images are captured by a drone.(Xuejin: proide the specific model.) Figure 3 shows the results of our experiments, where we project roof contours on the overhead image based on the estimated camera poses. We can see that our method is able to accurately align the over-head images captured at different positions with various buildings. The buildings are irregular, consisting parts of various heights.

The average time of register an 1300×600 overhead image to a 3D point cloud containing about $300K$ 3D points is about 8 minutes. (Xuejin: Check the numbers.) In the future, we intend to implement a faster version of the algorithm using parallel computing. Another future direction is camera tracking of a video sequence taking from overhead. This is useful for many drone-based applications.

REFERENCES

- C. Arth, C. Pirschheim, J. Ventura, D. Schmalstieg, and V. Lepetit. 2015. Instant Outdoor Localization and SLAM Initialization from 2.5D Maps. *IEEE Transactions on Visualization and Computer Graphics* 21, 11 (Nov 2015), 1309–1318. DOI: <https://doi.org/10.1109/TVCG.2015.2459772>

- M.-Y. Liu, C.O. Tuzel, A.N. Veeraraghavan, and R. Chellappa. 2010. Fast Directional Chamfer Matching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <http://www.merl.com/publications/TR2010-045>
- K. Ni, N. Armstrong-Crews, and S. Sawyer. 2013. Geo-registering 3D point clouds to 2D maps with scan matching and the Hough Transform. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. 1864–1868. DOI : <https://doi.org/10.1109/ICASSP.2013.6637976>
- X. Zhang, G. Agam, and X. Chen. 2014. Alignment of 3D Building Models with Satellite Images Using Extended Chamfer Matching. In *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 746–753. DOI : <https://doi.org/10.1109/CVPRW.2014.115>

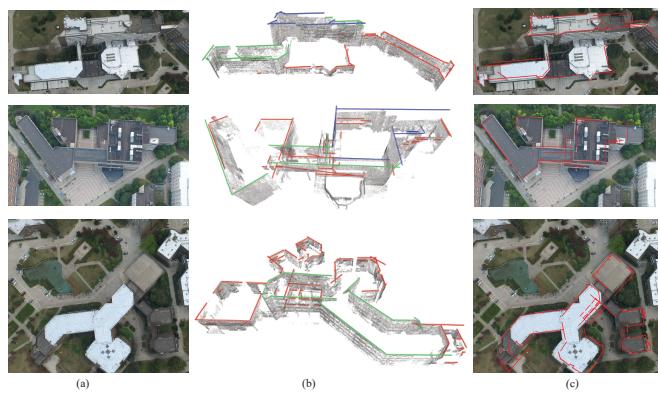


Figure 3: (a) The low-altitude overhead images. (b) Building point cloud with contours we extract. (c) Results of our experiments. We project the contours on the overhead image to show the results. (Xuejin: Why not use different colors for contours like in (b)?)