Accurate 3D Railway Track Extraction from Aerial Images

Dong Wei[a,1], Xiaotong Li[a,1], Yongjun Zhang[a,*], Chang Li[b] and Ziqian Huang[a,1]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430072, P.R.China
[c] College of Urban and Environmental Science, Central China Normal University, Wuhan, 430072, P.R.China

ARTICLE INFO

Keywords:
3D line segments
line clustering
line triangulation
3D line evaluation

ABSTRACT

Three-dimensional (3D) lines require further enhancement in both clustering and triangulation. Line clustering assigns multiple image lines to a single 3D line to eliminate redundant 3D lines. Currently, it depends on the fixed and empirical parameter. However, a loose parameter could lead to over-clustering, while a strict one may cause redundant 3D lines. Due to the absence of the ground truth, the assessment of line clustering remains unexplored. Additionally, 3D line triangulation, which determines the 3D line segment in object space, is prone to failure due to its sensitivity to positional and camera errors.

This paper aims to improve the clustering and triangulation of 3D lines and to offer a reliable evaluation method. (1) To achieve accurate clustering, we introduce a probability model, which uses the prior error of the structure from the motion, to determine adaptive thresholds;

1. Introduction

Currently, the length of the railway has exceeded 1.3 million kilometers on the earth, for which the maintenance and development of railways have a significant impact on safe operations. As the preliminary stage of extracting 3D railway track (RT) accurately and efficiently, to support engineering design, monitor construction quality, and ensure operational safety, has become one of the basic components in the maintenance of existing railways.

The extraction of RT can be achieved by real-time kinematics, LiDAR, and multiple images. The real-time kinematic is generally mounted on a railway measurement vehicle and obtains the RT by moving along the rail track. In general, it has a satisfactory accuracy while requiring operations on the track, thus demanding the cooperation of railway departments, and there are issues related to both safety and efficiency. LiDAR sensors can be mounted on a drone, which is more convenient and secure than real-time kinematic. Because a further process, like point segmentation or classification, is required for RT extraction, the drone must maintain a low flight altitude to satisfy the standards of the point-cloud density, which would impact the efficiency. A drone with cameras can capture aerial images efficiently with a safe distance from the railway area. But RT extraction is challenging in aerial images: (1) The dense points reconstructed with aerial images are inaccurate around the railway track because of the occlusion and matching problems caused by the parallax variation. (2) Joining image semantics to obtain RT might be workable; However, how to detect the semantics of RT accurately and completely in aerial images remains to be studied.

If we reconstruct the dense points from multiple aerial images and detect the RT from point clouds, the overall method of finding RT is similar to deal with the point clouds that obtained from mobile laser scanning (MLS) or airborne laser scanning (ALS). Generally, the RT can be detected with semantic segmentation, while the significant noise, inaccurate edge localization, and large density variations of point clouds bring about great challenges to the robust semantic segmentation. Thus, most general segmentation algorithm cannot be used directly in RT segmentation; instead, the carefully designed geometric priors was used to guide the segmentation and the grouping of RT: such as constructing the shape features and density data on the basis of railway bed extraction. However, these methods relies heavily on the quality and density of the point cloud, thus requiring the drone to maintain a low flight path to improve point cloud quality and reduce the processing range. Compared with point clouds, images contains rich semantic informations. Thus, several studies exploited the deep learning method that design the network for training and detect the RT from aerial images, which demonstrated the effectiveness of deep learning technology in RT extraction.

*Corresponding author
weidong@whu.edu.cn (D. Wei); zhangyj@whu.edu.cn (Y. Zhang); lichang@ccnu.edu.cn (C. Li)
[1] Co-first authors.

Moreover, the deep learning method relies heavily on training samples and considering the texture of railway regions varies greatly across the world, it may require an increased number of training samples to obtain a more generalizable detection network. In addition, these methods just deal with single frame and lacked the strategy of processing multiple aerial images.

This paper propose the accurate RT extraction for multiple aerial images, which fully exploits the contexture and geometry informations across multiple images. Compared to LiDAR based methods, we can use more affordable imaging drones to conduct a efficient and safer railway aera maping than ALS drones or MLS equipments. Compared to the former image-based methods, we propsose the complete clustering and reconstruction strategies that obtain the accurate and non-redundant 3D RT from multiple aerial images; moreover, in the clustering of RT across multiple images, we use the deep features of existing network trianed from imagenet, rather than detecting the RT with a new network; thus, non pre-training, which generaly needs expensive samples, is required.

References