Automation in Construction 85 (2018) 76–95

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3D local feature descriptor

Yusheng Xu*, Sebastian Tuttas, Ludwig Hoegner, Uwe Stilla

Photogrammetry and Remote Sensing, Technische Universitaet Muenchen, Arcisstr. 21, 80333 Munich, Germany

ARTICLE INFO

Keywords:
Scaffolds
Photogrammetric point cloud
3D local feature descriptor
Classification
Object reconstruction

ABSTRACT

Scaffolds always act as disturbances when reconstructing the 3D scene of the construction site due to occlusions, similarities with buildings in color and height as well as their adjacent positions to wall surfaces. Since scaffolds are commonly utilized to assist the construction and maintenance of building structures, professionals can estimate the overall progress and temporal objects of construction projects by assessing the status or arrangement of the scaffolds. Its thin, repeating and complex structures also make it a valuable dataset for testing related algorithms and approaches for the reconstruction of 3D construction site scene. To this end, we present a data-driven workflow for the detection and reconstruction of scaffolding components, including tubes, toeboards, and decks, given a photogrammetric point cloud. Our workflow consists of two parts: one part concerns the strategy based on projection and methods of grouping and slicing planar surfaces for detecting and extracting points of scaffolds from the construction site. The other part relates to the point feature derivation using a novel 3D local feature descriptor LSSHOT, designed for extracting features in the classification of points. Specifically, our workflow is implemented by five major steps, including preprocessing of the point cloud, division of building facades, classification of points, geometric modeling and refinement of results. To evaluate our proposed descriptor, a series of simulated experiments using synthetic datasets is conducted via shape matching tests. A real application is also carried out to validate the feasibility and effectiveness of our workflow using the photogrammetric point cloud of a construction site. Results of simulated experiments reveal that our proposed descriptor outperforms the original SHOT descriptor in the simulated test, especially when dealing with point clouds having a large percentage of noise. Regarding the real application of reconstructing scaffolds, points of scaffolds are successfully detected, extracted, and reconstructed. For a facade having enough points, over 70% of the scaffolding elements are reconstructed. For the classification of points using LSSHOT descriptor and a random forest classifier, the accuracy of results for the points of two major scaffolding elements reaches more than 70% in our test examples.

1. Introduction

1.1. Motivation

In the fields of Architecture, Engineering and Construction/Facility Management (AEC/FM), the demand for efficient and accurate progress monitoring of construction sites has increased in recent decades, driven by popular specialized applications in work progress tracking, productivity improvement, quality control, security assurance, accident investigation, collaborative communications [1,2]. Conventional progress tracking approaches largely depend on the visual inspection and require extensive manual collections of data and analysis of various documents. Such progress monitoring methods therefore not only rely heavily on the personal skills and experiences of professionals but also are fairly time consuming. To solve this problem, much attention has been devoted to the automatic construction site monitoring via 2D imaging, photogrammetry, and Terrestrial Laser Scanning (TLS) [2–4]. Among all the solutions, the reconstruction of 3D scene (e.g., the reconstruction of as-built BIM) using point clouds is increasingly widely used [3,4]. However, raw datasets such as the generated point cloud usually contain numerous secondary and temporary objects, for example, scaffolding components, which are deemed counterproductive to the reconstruction work. Scaffolds, commonly formed by thin structures like pillars and boards, are located adjacent to the building surfaces and share some similarities with the main structure of the building in shape, color and height. Due to these properties, scaffolding components may cause difficulties in the reconstruction of building structures resorting to occlusions, similarities and disturbances. Thus, if

* Corresponding author.
E-mail address: yusheng.xu@tum.de (Y. Xu).

http://dx.doi.org/10.1016/j.autcon.2017.09.014
Received 11 February 2017; Received in revised form 20 June 2017; Accepted 13 September 2017
0926-5805/

the scaffolding components can be detected prior to the processing of main building structures, the reconstruction work can be more accurate and efficient. Since the scaffolds are the most commonly used structures assisting the construction and the maintenance of buildings and they are also part of the construction site, we can roughly estimate the aggregate scheduling for the construction project by reconstructing scaffolds in a construction site. Moreover, the reconstruction of scaffolds can help us to test the related methods and algorithms about the reconstruction of the construction site, considering the complex and typical structure of scaffolds.

In this work, we detect and reconstruct scaffolding components of photogrammetric point clouds generated by multi-view stereo matching of a construction site with a complex environment, in order to make a good preparation for the further reconstruction work of as-built BIM and provide auxiliary information for the monitoring of the construction process.

1.2. Related work

Previous work on the detection and reconstruction of scaffolds using point clouds is scarce. The related work mainly focuses on the reconstruction of as built BIM from point clouds [3–6] or the comparison between the point cloud, as-built and as-planed BIMs [7–9]. To some extent, we deem scaffolds as constituent parts of a construction site. As common artificial objects in a construction site, scaffolds have a representative structure being quite similar to some building structures, with a skeleton consisting of tubes and boards with orthogonal connections. Therefore, with respect to the detection and reconstruction of such objects, a wide variety of approaches and algorithms related to the reconstruction of as-built BIM as well as the recognition of objects from point clouds can be referred and used. In our preliminary investigations [10], positive results have been achieved, with methods stemming from the reconstruction of as-built BIM and object detection. Hence, an appropriate method for recognizing objects or primitive shape is worthy of research and can facilitate our current reconstruction work.

Numerous attempts and contributions have been done in the fields of shape or object detection and reconstruction from point clouds. For example, in Ref. [11], a RANSAC-based algorithm is proposed to provide a valid solution for extracting several types of shapes from an oriented point cloud (i.e., point cloud with oriented surface normals). By means of statistical analysis and persistent histogram features estimation, in Ref. [12], Rusu et al. obtained geometric shapes of objects in a household environment from the point cloud through semantic 3D object maps. Okorn et al. proposed an approach using the Hough transform to extract planar segments from the point cloud [13]. Moreover, in Ref. [1] the object recognition is performed based on a threshold on the ratio of the covered area to the entire surface of the object. Additionally, in Ref. [14], Rottensteiner reviewed the local supervised classifiers and statistical models for the object extraction from LiDAR points in urban areas. In Ref. [15], Niemeyer detected buildings from point clouds via integrating the random forest (RF) classifier into the conditional random fields framework. In Refs. [15] and [16], the extraction of planar and linear features from laser scanning data is conducted by utilizing the principal component analysis (PCA) and its variations and extensions. Based on the previous work, in Ref. [17], Polewski demonstrated that the local 3D feature descriptors and local supervised classifiers can be used to efficiently detect segments of fallen trees in LiDAR point clouds. Among all the mentioned ideas and approaches, the ones based on local surface features extracted by various 3D feature descriptors show a promising prospect, because the unique and representative features they extract are distinctive when recognizing different types of objects and they are commonly not affected by the scale, rotation, and translation factors [18]. Accordingly, 3D feature descriptors can play a vital role for the task of object recognition, especially in the condition of the 3D reconstruction owing to the complex structures and background.

A 3D local feature descriptor is a compact representations of points based on the characteristics of their neighborhood (i.e., the so-called support region) [19]. The state-of-the-art 3D local feature descriptors can be divided into three main categories [18,20,21]: the descriptors based on the spatial distribution histograms of points in the neighboring support, the descriptors based on the geometric signature of the point in the local surface and the descriptors featuring a hybrid structure between spatial distribution histogram and geometric signature. In the first category, the descriptor usually defines a local reference frame or axis (LRF or LRA) for a key point, through which a 3D support region is divided into a number of bins in accordance with the LRF or LRA. Here, the key point is the one whose features needs to be extracted and normally with rich information content. By accumulating the spatial distribution information of points in the 3D support region into these bins, the histograms of descriptor are encoded. Some commonly used descriptors, for example, Spin Image (SI) [22] and its variant Tri-spin-image [23], 3D Tensor [24], rotational projection statistics (RoPS) [25], and 3D shape context (3DSC) [26] as well as its variations such as unique shape context (USC) [19] and Cylindrical-3DSC [17], all belong to this category. In the second category, the histogram of the descriptor is generated by encoding descriptive geometric attributes (e.g., normals or curvatures of points) of the key point and its surrounding points in the 3D support region. For descriptors of this category, point feature histograms (PFH) as well as the improved-efficiency version of fast point feature histograms (FPFH) [27], THRIFT [28], local surface patch (LSP) [29] and radius-based surface descriptor (RSD) [30] are regarded as representatives. In the last category, the descriptor has a hybrid structure combining the spatial distribution histograms and geometric signatures. For instance, signature of histogram of orientations (SHOT) [19] is a hybrid descriptor, encoding the histograms of the surface normal orientations in conjunction with different spatial locations of points in a spherical support region. Both the spatial location information of points and the local histograms encoding the angles of normal vector of points are combined in the overall histogram of SHOT.

For most of the aforementioned local feature descriptors, their performance exceedingly relies on the definition of the LRF or LFA [20] and the selection of the support region [31,32], which can significantly affect the robustness and accuracy of the descriptor [25]. However, in some cases using photogrammetric points, outliers and noise resulting from stereo matching normally have negative impacts on the definition of LRF or LRA, directly influencing the performance of the descriptor. Moreover, the shape of support region defining a respective neighborhood encapsulating all considered 3D points is also crucial to the representation of features obtained from the descriptor. Thus, developing a 3D local feature descriptor with a robust LRF and specific support region could be a possible solution for achieving improvement of accuracy in certain applications.

1.3. Objectives and contributions

In this paper, our objective includes two parts: one is to detect the points belonging to scaffolds from the photogrammetric point cloud of a construction site, while the other one is to reconstruct scaffolding elements from the detected points. As shown in Fig. 1a and b, the point cloud is generated from multi-view images. Improved approaches for the detection and reconstruction are given by extending a previous work [10] and put forward a detailed procedure of the recognition and reconstruction of fundamental elements of scaffolds: tubes, toeboards and decks, exhibited in Fig. 1c. Scientific contributions are as follows:

- A framework for the detection and reconstruction of geometric objects in the construction site: A framework of detecting and reconstructing scaffolds is proposed and validated, which is originally designed for the complex structures formed by multiple linear objects, with a success rate of around 65% in experiments of a construction site.