Advanced Engineering Informatics 38 (2018) 569–580

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Real-time collaborative reconstruction of digital building models with mobile devices

Steffen Franz*, Robert Irmler, Uwe Rüppel

Technische Universität Darmstadt, Institute of Numerical Methods and Informatics in Civil Engineering, Franziska-Braun-Str. 7, 64287 Darmstadt, Germany

ARTICLE INFO

Keywords:
Scan-to-BIM
Collaboration
Mobile devices
Google Tango
Building Information Modeling
Industry Foundation Classes

ABSTRACT

Many software-based building processes require digital building models. Since the building stock does not have sufficient data in this regard, the demand for Scan-to-BIM processes is increasing. In this paper we present a system for the reconstruction of 'as-built' BIM content of house interiors based on the Google Tango technology. The strength of our approach is the use of low-cost mobile scanning devices and a client-server system that allows for a real-time collaborative scanning and reconstruction of indoor scenes. We developed a server application that continuously aggregates scan data of multiple scanning devices (clients) and applies the data stream to a real-time post-processing pipeline to reconstruct rooms, walls, doors and windows. The reconstruction result is then distributed to all clients, where it is visualized in real time. The collaborative workflow and real-time data processing make our system especially useful in situations that are time-critical and require concurrent collection and processing of data. One of our targeted use cases therefore is the model generation for crime scene documentation. The effectiveness of our approach was demonstrated on three test sites. Our results compare well to other state-of-art methods regarding the reconstruction of walls, but they also revealed potential for improvement regarding the detection of doors and windows in occluded and cluttered environments.

1. Introduction

The comprehensive digital representation of a building provides a great source for computational analysis, optimization and building operation. In recent years, the concepts and technologies of Building Information Modeling (BIM) have been adopted by a great share of the Architecture, Engineering and Construction (AEC) industry worldwide. Building Information Models are mostly created during the design and costing phases of construction projects. Beyond that, there is also an increasing demand for the digital reconstruction of the existing building stock ('Scan-to-BIM'). Typical applications that require digital 'as-built' models are for example, energy simulations and facility management. In addition, the reconstruction of semantically rich building models opens up a wide range of new application areas, ranging from the preservation of architectural knowledge [1] to the forensic documentation of fire scenes [2]. Over the past two decades, the technologies for the digital reconstruction of our environment advanced rapidly. In this context, the robust reconstruction of complete indoor scenes represents a particularly challenging problem. Automated approaches usually employ extensive and costly hardware setups and often involve time consuming scanning and evaluation steps [3,4]. Another common drawback of such solutions is the separation of data collection and processing. Without immediate feedback, errors induced during the scanning process can only be recognized after the fact and might require a time consuming and costly rescanning. Furthermore, previous work in the field of indoor reconstruction often focused solely on achieving a high accuracy of scans and reconstructed geometry. Time aspects, hardware costs and computation resources are usually of subordinate importance. Because of that, applications that are time-critical or require the concurrent collection and processing of data can hardly profit from these developments. In this context, current developments in the field of mobile devices present promising alternatives. Modern devices, like Google Tango-enabled smartphones, combine multi-modal sensor technologies, including depth sensing, with significant computing power. In addition they are affordable and easy to handle. We therefore propose a framework for collaborative indoor reconstruction based on mobile depth sensing devices to speed up the Scan-to-BIM process at low costs. Our framework is based on a client-server architecture that allows for the collection and distribution of scan data. Multiple mobile scanning devices are used to collect scan data of an indoor scene. The devices continuously send their data to a server application which aggregates it and applies the resulting data

* Corresponding author.
E-mail address: sekretariat@iib.tu-darmstadt.de (S. Franz).

https://doi.org/10.1016/j.aei.2018.08.012
Received 7 December 2017; Received in revised form 8 May 2018; Accepted 15 August 2018

stream to a reconstruction pipeline. The pipeline constantly performs a geometric reconstruction based on the aggregated scan data and detects spaces, walls and openings. The reconstruction result is immediately made available to all connected clients, thus enabling real-time collaborative reconstruction of digital building models. Finally, the result can be exported as an IFC model (Industry Foundation Classes) to allow subsequent processing.

1.1. Use-cases for real-time collaborative reconstruction

Apart from speeding up the scanning process by using multiple devices, there are a variety of use-cases that can particularly profit from a collaborative and real-time reconstruction procedure. One example is the digital documentation of crime scenes. Current research on crime scene documentation shows, that this field can greatly benefit from a digital model of the site. For example [5,6], utilized a 3D crime scene reconstruction for documentation purposes and emphasis the resulting benefits for presentations and reenactments in court, testing theories or examining witness statements. Typically, stationary and mobile laser scanners as well as photogrammetry processes are used to record a scene. Following the data collection, the scan data needs to be transferred to a computing device for processing. This usually requires specialized software and skilled personal. Meanwhile, the evidence on-site is mainly documented in form of pictures, videos, sketches, notes and voice-recordings. The spatial and semantic relations between these distributed pieces of information have to be generated manually. For example, according to the National Fire Protection Association guidelines [7], photographs of fire scenes have to be documented in a floor plan sketch along with their name, location and orientation to ensure a coherent and intelligible preservation of evidence. Because crime and fire scene documentations are highly time-critical tasks, they could particularly profit from a collaborative and real-time digital reconstruction of the investigated scene. A mobile device that provides a live reconstruction of the site would allow for an immediate annotation and localization of evidence within the scene while the user is still on-site. Photos taken by the mobile device or a connected camera would automatically be localized within the digital model. Mandatory sketches could then automatically be extracted form the digital model. The parallel use of multiple, connected scanning devices could significantly speed up the documentation process. Moreover, by instantly sharing information over a network, it would further be possible to have external experts participating 'live' in the documentation process. Other application areas that could profit from a real-time, mobile and collaborative digital reconstruction of building structures are renovation and construction work targeting the building stock. Architects, engineers and craftsmen are often required to perform on-site measurements of dimensions to calculate materials and costs correctly. This is usually achieved by a manual survey of the site. Although modern handheld distometers measure distances and spaces fast an accurate, their measurements still need to be processed to acquire usable results e.g., floor plan, masses. Especially in large or complex environments, this method quickly becomes time consuming and prone to errors. In this context, the utilization of low-cost mobile devices and the instant availability of a digital building model could be of great use, e.g., regarding mass calculations or progress tracking on-site.

1.2. Outline

This paper consists of six sections including the introduction. Section 2 presents a brief review of methods targeting the reconstruction of 'as-built' building models. Section 3 discusses requirements for a collaborative reconstruction process and presents our implemented framework based on the Google Tango technology. In Section 4 we explain in detail the real-time processing of the scan data and describe our approach on detecting building structures like spaces, walls, doors and windows. We evaluated our framework on several test sites. The results are presented in Section 5. Finally, Section 6 summarizes conclusions and outlines starting points for further research.

2. Related work

Over the past two decades, a lot of effort went into exploring new technologies for the digital reconstruction of our environment. The necessity for the digital reconstruction of cities and buildings arises from various application areas e.g., urban planning, pollution forecasting, energy efficiency simulation or disaster management.

Spatial data at city scale is usually collected via airborne LiDAR systems [8–10]. To improve data quality, the fusion of different data sources like images and GIS data has also been investigated [11]. For the outside reconstruction of single buildings, stationary laser scanners [12] and image-based techniques [13–15] have been proven to be suitable. The robust reconstruction of complete indoor scenes however, remains a particularly challenging problem. Indoor environments often exhibit high levels of clutter and occlusion which can lead to noise, artifacts and missing data. Reliable indoor reconstruction methods therefore have to be particularly tolerant to missing data, occlusion and highly reflecting or transparent surfaces like mirrors and windows. Prior work on this subject shows a great variety of different approaches each of which displays strengths and weaknesses. Many indoor reconstruction systems require extensive hardware setups [4,16,3,17,18], rely on user interaction [19,2,20] or only generate meshed geometries, i.e., do not identify individual building elements and therefore lack semantic depth [21,22]. In particular, the detection of windows is an especially challenging task. Due to their transparency, windows are shapeless in 3D space and lack discriminative power such as color and texture for image feature descriptors. Approaches on the detection and reconstruction of windows include computer vision and machine learning [17,23,24].

In the past, many of the mentioned approaches have been based on stationary laser scanning or image processing methods (e.g., [4,25,18,26,27]). In recent years, with the availability of consumer depth sensor devices as well as powerful mobile devices, an additional research and application field emerged. In [28,29] the authors present smartphone applications for generating 2D floor plans and 3D models under a Manhattan-world assumption [30] based on a combination of internal sensor data and photogrammetric modeling. Other research investigated consumer depth cameras like the Microsoft Kinect or the Google Tango as a cost-effective way to gather 3D measurements of the environment [2,31,32]. However, these technologies do not themselves solve the underlying problems to automatically extract higher-level models (semantics) from the acquired raw data. Most work on scene reconstruction with consumer depth cameras focused mainly on reconstructing a detailed mesh of the scanned environment [21,22], rather than generating a semantically enriched building model composed of detected building elements (e.g. spaces, walls, windows). Furthermore, time aspects, hardware costs and computation resources have usually been of subordinate importance.

3. A framework for real-time collaborative reconstruction of digital building models with Google Tango

This section introduces a framework for the real-time collaborative reconstruction of digital building models based on Google Tango-enabled mobile devices. The Google Tango technology was chosen because these devices have the advantage of being all-in-one solutions that combine depth perception, indoor navigation, processing power, and user interface. Other mobile devices capable of depth perception and visual odometry are, for example, the Intel RealSense DepthCamera [33], the Occipital Structure Sensor [34] and the ZED Stereo Camera [35]. These compact sensors can easily be combined with mobile computers like tablets and laptops, providing an alternative to Google Tango.