A Smart-Site-Survey System using Image-based 3D Metric Reconstruction and Interactive Panorama Visualization

Sha Yu, Kevin McGuinness, Patricia Moore, David Azcona, Noel O'Connor
SFI Insight Centre for Data Analytics
Dublin City University, Ireland
sha.yu@insight-centre.org

ABSTRACT

This work presents a so-called Smart Site Survey (SSS) system that provides an efficient, web-based platform for virtual inspection of remote sites with absolute 3D metrics. Traditional manual surveying requires sending surveyors and specialised measuring tools to the targeted scene, which takes time and requires significant human resource, and often includes human error. The proposed system provides an automated site survey tool. Sample indoor scenes including offices, storage rooms, and laboratory are used for testing purposes, and highly precise virtual scenes are restored, with the measurement accuracy of 1%, i.e. an error ±1.5cm to a 150cm length. This is comparable or superior to existing works or commercial products.

CCS CONCEPTS

• **Computing methodologies** → **Virtual reality**; **Reconstruction**; **Scene understanding**.

KEYWORDS

Virtual site survey, 3D reconstruction, panorama visualisation.

ACM Reference Format:
Sha Yu, Kevin McGuinness, Patricia Moore, David Azcona, Noel O'Connor. 2020. A Smart-Site-Survey System using Image-based 3D Metric Reconstruction and Interactive Panorama Visualization. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The last decade has seen ever-increasing demands of digitally produced virtual versions of real-world objects or scenes. Various applications range from game development [5], to virtual touring [2, 4]. This demo describes a novel virtual-scene modelling and visualisation system that can accurately survey the spatial dimensions and layout of indoor scenes, with a special focus on industrial sites linked with logistics and supply chain services. For example, storage rooms on a (maritime) transport vessel. Traditional methods for conducting site surveys of such spaces rely on manual measurements, which are time consuming and prone to human errors. Due to constraints on site access, sending or setting up specialised measurement devices can be difficult. We are particularly interested in scenes where the space layout can require frequent updating where the manual site surveying becomes particularly inefficient.

Existing techniques for 3D-scene modelling generally fall into one of two categories - active or passive [1, 3]. Active approaches utilise 3D sensing of objects/scenes that use laser technology, structured light, or large camera arrays. These technologies require trained personnel, stable environments, and can be prohibitively expensive. Passive approaches typically rely on the processing of images for model generation, and are less constrained than active approaches. Such approaches also do not require specialised equipment or trained personnel, so the nature of the environment's constraints on capture are less crucial.

This work decides to choose the passive technique based 3D reconstruction. In our case, we are concerned with scenes that are relatively small-scaled and/or cluttered, have relatively poor illumination. These situations will present challenges for image-driven 3D reconstruction: constrained by the space, the camera may take in only a small part of the environment at a given time; The weak illumination can affect the image quality. Although current state-of-the-art algorithms have demonstrated successes at different aspects, holes or incomplete surfaces are still commonly observed in final 3D outputs. This is especially the case for indoor scenes, as reflected by well-known 3D reconstruction or multi-view stereo benchmarks, e.g. Tanks and Temples – Intel 2017 SIGGRAPH. In this work, we propose to use a panorama scene model as our visualisation interface, rather than directly displaying an incomplete 3D model. The 3D point cloud resulted from the image-based reconstruction process will be projected onto the panorama field with absolute 3D metrics implicitly encoded in this representation. Existing solutions rely on camera arrays, pre-calibrated cameras, or IMU sensors for the purpose of metric 3D reconstruction. These instruments are not feasible for our application: First, dynamic environments are involved with modern logistics and transport services, a high-throughput pipeline of virtual scene modelling is crucial; Secondly, this project considers the reconstruction of mobile sites during transportation, where the external motion of the scene might add considerable noise to IMU readings.

We present an end-to-end pipeline from capture to visualisation. Our main contributions include: First, a high-throughput site survey system is developed, which is particularly beneficial when targeted scenarios involve dynamic or frequently updated space plans. Second, accurate 3D metrics are realised with a pure image processing strategy, which is validated in environmental conditions that may invalidate commonly adopted sensors or equipments.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 SMART SITE SURVEY SYSTEM

The proposed system consists of three functional layers.

Layer 1: 2D-data acquisition and uploading. The only hardware we need is a regular smart-phone with a web-browser and a camera capable of 2D image capture. To facilitate the later 3D modelling and panorama building process, the input 2D data requires a set of still images, captured from different viewpoints and with high degree of overlap, and a video acquired by slowly rotating the (phone) camera to capture as much of the scene as possible.

Layer 2: backend computer vision processes, including a standard 3D modelling process using modern structure from motion (SfM) and multi-view stereo vision techniques, image stitching based panorama generation, reference recognition based metric estimation, and 3D-to-2D mapping from 3D space to panorama view. The only requirement for our approach is QR codes (i.e. references) placed in the scene, due to the factor that images-based 3D modelling requires at least one absolute measurement to achieve 3D metric reconstruction. For remote sites QR codes can be easily printed and placed in the scene by non-expert users. In order to automatedly segment the reference object, machine-learning based pattern detection, and projective transformation based template matching are sequentially applied to locate the reference region.

Notice that in our approach, features are projected from the 3D point cloud onto the panorama field (see illustration in Figure 1). Panorama stitching usually assumes a different camera-motion rule to SfM in 3D modelling — for the former, only rotation motions are assumed for the camera. Because of this, the 3D point cloud has no direct path to be projected onto the panorama space. We introduce a two-step process as follows: firstly we project the 3D points onto the (video) images, and then map the 2D features to the panorama field, according to two camera spaces respectively.

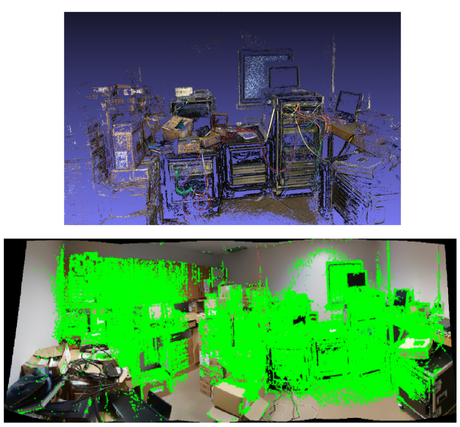

Figure 1: The 3D features are projected from the point cloud space (top) onto the panorama field (bottom).

Layer 3: the web-service interface. This has two roles: 1) Providing an entry point for smart-phone users to upload 2D data. Specifically, a URL address that links to the interface is announced once the computer server gets started. Each 2D-data upload event will trigger the backend server to run a new procedure for scene modelling. The result produced by the server will be fed to the web service. 2) Offering the visualisation and interaction platform. Terminal users will be able to browse pre-reconstructed scenes, with interactive actions to zoom in/out or shift the panorama view. By clicking-and-dragging lines into the virtual scene, the corresponding 3D metrics will be displayed.

Figure 2 illustrates the web pages that respectively serve virtual room selection, and interactive survey. The backend server of the system is invisible to users, where the 3D modelling and the associated image processing steps are hidden. In the demo we will demonstrate the 3D reconstruction and panorama mapping process and the use of the visualisation/measurement interface.

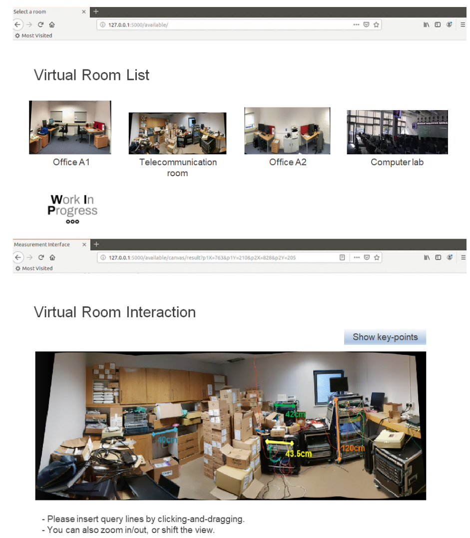

Figure 2: Illustration of the interaction interfaces.

In this system, the computer vision functions requiring most computation are coded in C++, with supports from open-source libraries of Colmap, OpenCV, OpenCV-Python, Eigen, etc. The web-service platform is developed with libraries or tools including Flask, Bootstrap, Celery, Redis, and SQLite. The server is executed on a computer with an AMD Ryzen5-2600, running at 3.90GHz with 6 cores and 16GB of RAM, and with a Nvidia (GeForce Rtx 2080) 11GB graphics card. Depending on the number of 2D images being captured (ranging from 30 to 50), the running time for each whole process is roughly between 0.5 to 2 hours. For practical usage, a more powerful server or GPU cloud is able to further reduce the operation time of the procedure.

The developed system achieved comparable or better performance, in comparison with recently published products that also focus on metric 3D reconstruction, or scene-understanding based metric estimation. The commercial product Matterport reports a measuring accuracy also within 1%, however it requires compatible cameras. The AR-measuring apps (developed by Google, Apple) demonstrates a fluctuating estimation accuracy: the error approx. ±1.5cm is achieved for real scales in the range of 5cm to 20 cm; For larger devices with dimensions about 1 to 2 metres, the deviations of the automated metrics may exceed ±10cm.

ACKNOWLEDGMENTS

This project is co-funded by Enterprise Ireland, Science Foundation Ireland Insight Centre for Data Analytics, and Panalpina World Transport (Ireland) under the Innovation Partnership Programme.

REFERENCES

[1] Irene Aicardi, Filiberto Chiabrando, Andrea Maria Lingua, and Francesca Noardo. 2018. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. *Journal of Cultural Heritage* 32 (2018), 257–266.

[2] Belen Jiménez Fernández-Palacios, Daniele Morabito, and Fabio Remondino. 2017. Access to complex reality-based 3D models using virtual reality solutions. *Journal of cultural heritage* 23 (2017), 40–48.

[3] Richard Hartley and Andrew Zisserman. 2003. *Multiple view geometry in computer vision.* Cambridge university press.

[4] Anestis Koutsoudis, Fotis Arnaoutoglou, and Christodoulos Chamzas. 2007. On 3D reconstruction of the old city of Xanthi. A minimum budget approach to virtual touring based on photogrammetry. *Journal of Cultural Heritage* 8, 1 (2007), 26–31.

[5] Eric Turner, Peter Cheng, and Avideh Zakhor. 2014. Fast, automated, scalable generation of textured 3d models of indoor environments. *IEEE Journal of Selected Topics in Signal Processing* 9, 3 (2014), 409–421.

[6] Yanxing Zhang, Lei Li, and Beibei Liu. 2019. The Discussion on Interior Design Mode Based on 3D Virtual Vision Technology. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 23, 3 (2019), 390–395.