

Crowdsourcing of building interior models

Julian Rosser
Horizon Doctoral Training Centre
Nottingham Geospatial Institute
University of Nottingham
Nottingham, U.K.
psxjr@nottingham.ac.uk

Jeremy Morley
Nottingham Geospatial Institute
University of Nottingham
Nottingham, U.K.
jeremy.morley@nottingham.ac.uk

Mike Jackson
Nottingham Geospatial Institute
University of Nottingham
Nottingham, U.K.
mike.jackson@nottingham.ac.uk

Abstract

Indoor spatial data forms an important requirement underlying many ubiquitous computing applications. It gives context to users operating location-based applications, provides an important source of documentation of buildings and can be of value to computer systems where an understanding of environment is required. Unlike external geographic spaces, no centralised body or agency is charged with collecting or maintaining such information. This paper takes the position that models of building interiors can be volunteered by users of these spaces. The widespread deployment of mobile devices provides a potential tool that would allow rapid model capture and update. Here we discuss issues and requirements of a building interior modelling system together with a preliminary method for capture of this information. The nature of indoor data is inherently private; however these issues and legal considerations are not discussed in this paper.

Keywords: building modelling, indoor mapping, VGI, mobile.

1 Introduction

Digital modelling of buildings is required for a wide variety of purposes. As such, the topic has been approached in different ways from a variety of disciplines. Architecture, engineering and construction (AEC), navigation, positioning, robotics and emergency management are just some of the disciplines and application areas that have a stake in improving the digital capture, update, management and utilisation of building models. The term “building model” refers to digitally encoding useful information on both the geometric structure and the semantics of buildings. Included within the model is information pertaining to the interior of the building which is focused upon here.

One motivating application for deriving interior models of environments is their use in different positioning and navigation scenarios. For example, this information would be useful for general navigation and routing, first-responders in emergency situations and individuals with visual or physical disabilities. Locating people and objects with reference to an interior map is not an application restricted to large environments. Explicit spatial modelling of home environments is required if smart buildings and software agents are to understand a user’s context and activity. Indoor spatial models that hybridise geometric and symbolic information offer a promising approach for fulfilling the range of uses required for context-aware applications [1]. 3D building models can also be used for improving Augmented Reality (AR) systems which can benefit from provision of a prior initial scene model [12].

The need for digital encoding of indoor spaces has been identified as an important development for the infrastructure underlying many ubiquitous computing scenarios. The wide-range of applications demonstrates a need for standardised datasets which maybe generalized across the range of uses. But currently there is no ready source of this information.

Google has recently launched a service for businesses to upload their own floor-plans for display in Google Maps. Meanwhile ongoing research is looking at how to extend OpenStreetMap to indoor environments through extending the existing tagging scheme with a 3D ontology of building exteriors and interiors [4,5].

The capture and management of building interior data presents a different set of technical requirements and social challenges to those found when working with outdoor environments. To derive 3D virtual cities which model building exteriors on a wide scale, data from air- and space-borne sensors are often used in a semi-automated or fully automated workflow [6]. In an industrial context, 3D city model generation may be carried out by National Mapping Agencies (NMA’s) or companies with expertise in remotely sensed data processing. But, unlike exterior geographic spaces, no centralised body or agency is charged with the responsibility of collecting or maintaining the building interior details. To maximise the utility of these models they should be described in a suitable reference frame and standardised in a way that allows their integration with other spatial data.

The rest of the paper is structured as follows: Section 2 describes the high-level design issues surrounding a system to volunteer building interior models; Section 3 details requirements of building geometry capture; Section 4 outlines a preliminary data capture application; lastly, Section 5 concludes with a discussion and consideration of further work.

2 Background and concept

Crowdsourcing interior data for building models enables a building’s users to capture and manage data pertaining to these public and private spaces. The widespread availability of mobile computing devices incorporating cameras and other

sensors offers a possible tool for initially producing and maintaining interior building data. The successful application of crowdsourcing to geographic data capture has already been shown in projects such as OpenStreetMap and is in keeping with the notion that GIS should be resident-generated [14]. As such, it makes sense to frame the proposed research around a similar approach.

2.1 Motivation

Both public and private indoor spaces require digital modelling to fulfil the needs of applications outlined in Section 1. Providing that appropriate management tools for model data can be provided, the users of a building are in many ways the best sources of the information about it. They have access to the building's rooms, know about different levels of accessibility and understand roles fulfilled by particular areas. As the use of space within buildings can change, timely updates to the information can be provided. Furthermore, previous model versions may be retained. This update of the information provides a strong advantage over traditional top-down acquisition methods, such as laser scanners. Putting the user in control of the interior data is a primary motivator for this research.

It is acknowledged that in some circumstances users may provide incomplete or inaccurate information and this may be accidental or even deliberate. Thus any system for the curation of crowdsourced building models must be designed to cope with disparate and dynamic user input. As an example, an apartment block comprises many adjoining private homes. Each resident can model their own space but may not have direct access to or visuals of their neighbour's data. Behind the scenes, however, logical and topological checks, quality control and optimization of submitted data can be performed to help maximize the probability of useful contributions.

Imagery forms an important source of information in addition to geometric and semantic information on interior spaces. Not only can the imagery be used to texture building models but it can also provide relevant landmarks to users navigating a particular route. Utilisation of landmark information within electronic navigation aids has been found to have significant advantages over paper maps, enabling shorter navigation times and the lowering of mental and physical demand, and is particularly effective with older users [7]. Design guidelines on navigational services for pedestrians that state that landmarks should be used as the primary means of routing information [10].

As the technologies for positioning indoors with associated location-based services mature, the number of users capable of active or passive contribution will obviously increase. So we can already recognise that users of an interior modeling system might fulfil two roles: either as a contributor or an application user. However, an application user may also be fulfilling a secondary role of active/passive data update.

Issues of privacy are inherent in the modelling of interior spaces. Firstly, the use of video or image capture methods

requires careful consideration. The intention is that 3D geometric features are rendered with textures provided by the imagery; however, these may contain personal details that should not be made publically accessible. For example, the presence of particular people or valuable objects may be identifiable from the imagery. Furthermore, the model information itself can also be considered personal information.

3 Capture of building geometry

For the rest of this paper we mainly consider the acquisition of the geometric data of the building model. In this section, we outline what we consider to be important requirements for model construction in a crowd-sourcing context and why many 3D reconstruction approaches are unsuitable, namely that such a system should be available to common handsets and thereby allow frequent model update.

3.1 Commonly available devices

For high levels of metric accuracy, reconstructing the 3D geometry of indoor environments can be achieved using specialist data capture mechanisms. Terrestrial laser scanners are typically used, providing dense point clouds which can then be modelled in 3D [2,11]. However, the specialist nature of this equipment means its use is restricted and updates may be infrequent. Consumer-grade depth-sensing cameras (such as Microsoft's Kinect) are available but are not yet ubiquitous. For consumer modelling, high accuracy is of less importance and accessibility is critical.

3.2 Feature-less environments

The topic of 3D reconstruction is a long-standing topic within the computer vision research community. A common approach to generating a 3D point cloud from a sequence of image or video frames is to use a Structure-from-motion pipeline. This style of approach has been demonstrated on user-generated imagery [13] and indoor environments [3]. Similarly, Visual Simultaneous Localisation and Mapping (SLAM) techniques generally aim for online estimation of camera pose estimation and key feature locations. The PTAM framework [8], for example, is designed for AR usage and has been shown to work on a mobile phone [9]. However, both SLAM and Structure-from-motion approaches are not yet robust for all indoor scenarios. Large sections of interior walls either lack the texture required for feature detection and matching or require long model generation times.

3.3 Semi-automatic modelling and clutter

The constraints detailed in 3.1 and 3.2 indicate a need for a semi-automatic modelling process. And furthermore, indoor environments tend to be cluttered by furniture or other objects which occlude wall surfaces – the most desirable features for reconstruction. It could be argued that objects such as furniture should be included within interior models.

Recording the location of furniture would be beneficial for certain applications, such as navigation, particularly when visual senses are impaired as in an emergency response scenario. However, it is better to determine the true wall position and then populate the model with further information on objects contained within it. Although making the entire reconstruction process automatic is appealing, it is accepted that the final interaction to identify structure must have user input at some stage.

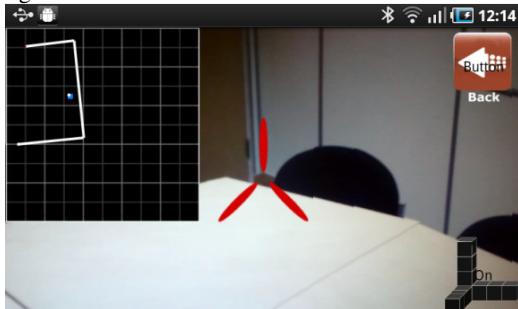
4 Preliminary data capture application

In this section we present ongoing work on the development of a system for in-field modeling of building interiors. The aim of the application is to incorporate a simple semi-automatic modelling process which may be used as a starting point for building data capture. The application has been implemented for Google Android and tested using the Samsung Galaxy Nexus phone.

4.1 Implementation

To capture the geometric structure of a room, application users fulfil an interaction similar to capturing a panoramic photograph - considered to be relatively familiar to many potential operators. Standing approximately in the centre of the space, the user points the device camera at key features located at ground level, namely floor-wall corners and points along the floor-wall intersection (see Figure 1). As the user rotates, denoting key features, a 2D plan view polygon is constructed.

Figure 1: Semi-automatic mobile room reconstruction.



For each point denoted by the user, the device orientation (azimuth, pitch and roll) is computed according to the device's accelerometer gravity vector, magnetometer and gyroscope. Two-dimensional coordinates of points with respect to the device are then estimated using trigonometry, assuming a fixed, user selected, device height across all points. During the capture process the user also collects at least one pair of corresponding ceiling-level and floor-level points from which the room height can be estimated, again using trigonometry. A point submission also triggers the capture of a photograph for association with the location. Finally, the model is geographically located by the user with translation, rotation and scaling actions against Google Maps imagery layer. The resulting points are extruded to form a 2.5D polygon.

This simple modeling approach contains a variety of sources of error based on several key assumptions: an accurate height measurement for the device, an unobstructed view of key features, a single-level floor and ceiling and a reasonably accurate orientation estimate. These issues contribute to irregularities apparent in the resultant 2D polygon. To compensate for them we regularize to a Manhattan-world where all key points are orthogonal to or parallel with each other using constrained least-squares.

4.2 Example model and discussion

An example model of a typical L-shape type office room within the University has been produced to illustrate the mapping process. Six points and images were collected. To illustrate the level of correction applied by constrained model fitting, Figure 2 shows both estimated feature locations together with the corrected floor-plan polygon. The model may be displayed as an extruded KML polygon in Google Earth (see Figure 3).

Figure 2: Estimated and corrected feature points (top-down perspective)

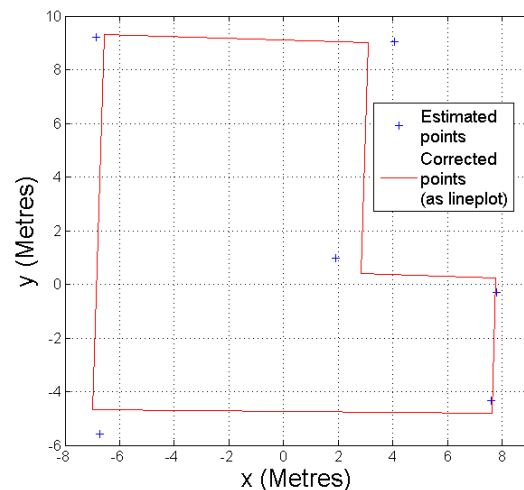
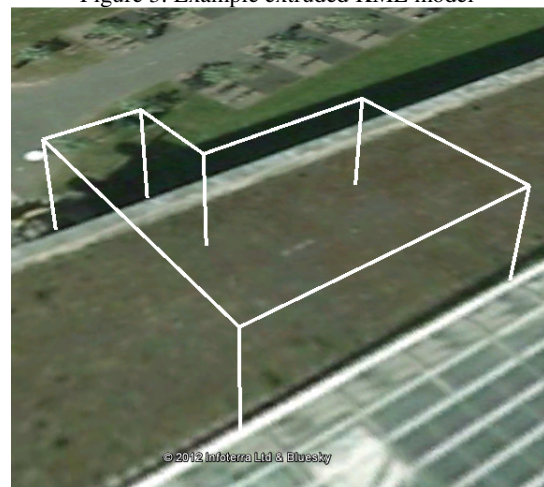


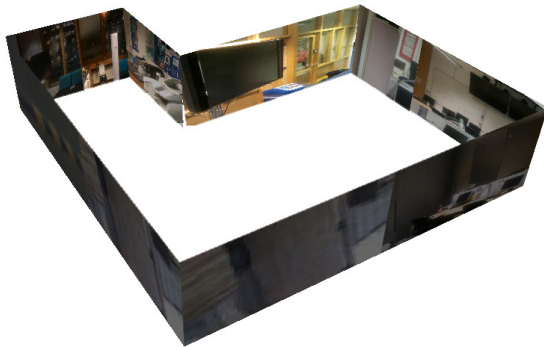
Figure 3: Example extruded KML model



Base Map Source: InfoTerra Ltd & Bluesky 2011.

Work is ongoing to finalise the application for user testing and release with several modifications envisaged for improvement. Of most importance is the facility to capture non-Manhattan world environments. An approach for this is to model each wall with many vertices followed by fitting of walls to each other. Secondly, the simple modelling approach described is most appropriate for relatively small spaces and where the user can stand centrally to view all points. A piecemeal approach of measuring walls from different locations would improve accuracy and increase the range of scenarios the tool is suitable for. Lastly, texturing of the model with captured imagery is planned. Room images may be manually edited and positioned using desktop software for texturing (see Figure 4). However, ideally the textures for each 3D face would be extracted and projected automatically. Initial tests indicate that some degree of user interaction would be required for meaningful projected texturing.

Figure 4: Manually textured model



5 Concluding remarks and future work

This paper presented a design overview, requirements for reconstructing interior geometry and a preliminary data capture application for input of volunteered building information.

While much research effort is directed on 3D reconstruction less attention has been given towards the management, conflation and update of these models. For example, in our data capture application, the initial model scale is dependent on an accurate user selection of device height. The model scale is also changed during the process of geo-referencing and hence the overall representation will vary with different users and levels of base mapping available. Absolute scaling is a common problem with many reconstruction methods that rely on geometric cues or arbitrary estimations of movement.

Furthermore, volunteered building models will likely to grow in an incremental and disparate fashion, but still require efficient storage with appropriate structural and topological constraints. Accurate user placement of interior models against a base map can be difficult and although aerial imagery or building footprints provide some spatial reference, users need to model spaces where there are few existing geographic features to refer to. We envisage employment of a rule-based engine to help optimise and enforce constraints,

leaving the possibility of sourcing data from a variety of sources and methods. It is also important to note that information capture is required not only initially - when no pre-existing model is available - but also when change to the geometry, texture or semantics of the space is apparent. Therefore, consideration of the infrastructure requirements for the combining and editing of these models is needed and will form an important part of further work.

References

- [1] Imad Afyouni, Cyril Ray, and Christophe Claramunt, Spatial Models for Indoor & Context-Aware Navigation Systems : A Survey. In *Journal for Spatial Information Science*. 2011.
- [2] Angela Budroni and Jan Böhm, Automatic 3d Modelling Of Indoor Manhattan-World Scenes From Laser Data, *Archives, XXXVIII* 2010.
- [3] Yasutaka Furukawa, Brian Curless, Steven M. Seitz, Richard Szeliski. Reconstructing Building Interiors from Images. *ICCV*. 2009.
- [4] Marcus Goetz M., Alexander Zipf. Extending OpenStreetMap to Indoor Environments: Bringing Volunteered Geographic Information to the Next Level. In: Rumor, M., Zlatanova, S., LeDoux, H. (eds.) *Urban and Regional Data Management: Udms Annual 2011*: Delft, The Netherlands. p. 47-58. 2011.
- [5] Marcus Goetz. Alexander Zipf. Towards Defining a Framework for the Automatic Derivation of 3D CityGML Models from Volunteered Geographic Information. In *International Journal of 3-D Information Modeling*. 1(2). IGI-Global. 2012.
- [6] N. Haala and M. Kada, An update on automatic 3D building reconstruction, In *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, Nov. pp. 570-580. 2010.
- [7] J Goodman, S A Brewster, and P Gray, How can we best use landmarks to support older people in navigation ?, *Behaviour*, 24. 3 - 20. 2005.
- [8] Georg Klein and David Murray, Parallel Tracking and Mapping for Small AR Workspaces, In *6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE, 1-10. 2007.
- [9] Georg Klein and David Murray, Parallel Tracking and Mapping on a camera phone, In *8th IEEE International Symposium on Mixed and Augmented Reality*, IEEE, 83-86. 2009.
- [10] Andrew J May, Tracy Ross, and Steven H Bayer, Pedestrian navigation aids : information requirements and design implications, *Transport*. 331-338. 2003.

[11] Brian E. Okorn, Xuehan Xiong, Burcu Akinci, and Daniel Huber, Toward Automated Modeling of Floor Plans. *Proceedings of the Symposium on 3D Data Processing, Visualization and Transmission*, May, 2010.

[12] Gerhard Reitmayr and Tom Drummond, Going out: robust model-based tracking for outdoor augmented reality, *IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE. 109-118. 2006.

[13] Noah Snavely, Steven M. Seitz, Richard Szeliski, Photo tourism: Exploring photo collections in 3D. *ACM Transactions on Graphics (SIGGRAPH Proceedings)*, 25(3), 835-846. 2006.

[14] E Talen, Constructing neighborhoods from the bottom up: the case for resident-enerated GIS, *Environment and Planning B: Planning and Design*, 26 533-554. 1999.