



City Research Online

City St George's, University of London

Citation: Brujic-Okretic, V., Gatzidis, C., Liarakapis, F. & Baker, S. (2008). Towards Rapid Generation and Visualisation of Large 3D Urban Landscapes for Mobile Device Navigation. Paper presented at the IEEE Virtual Reality Workshop on Virtual Cityscapes: Key Research Issues in Modeling Large-Scale Immersive Urban Environments, 8 - 9 Aug 2008, Reno, Nevada.

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/204/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Towards Rapid Generation and Visualisation of Large 3D Urban Landscapes for Mobile Device Navigation

Christos Gatzidis¹
giCentre, Department of
Information Science,
City University

Vesna Brujic-Okretic²
giCentre, Department of
Information Science,
City University

Fotis Liarokapis³
Department of Creative
Computing, Coventry
University

Stuart Baker⁴
Alcatel-Lucent Telecom
Limited UK

ABSTRACT

In this paper a procedural 3D modelling solution for mobile devices is presented based on scripting algorithms allowing for both the automatic and also semi-automatic creation of photorealistic quality virtual urban content. The combination of aerial images, GIS data, 2D ground maps and terrestrial photographs as input data coupled with a user-friendly customized interface permits the automatic and interactive generation of large-scale, accurate, georeferenced and fully-textured 3D virtual city content, content that can be specially optimized for use with mobile devices but also with navigational tasks in mind. Furthermore, a user-centred mobile virtual reality (VR) visualisation and interaction tool operating on PDAs (Personal Digital Assistants) for pedestrian navigation is also discussed. Via this engine, the import and display of various navigational file formats (2D and 3D) is supported, including a comprehensive front-end user-friendly graphical user interface providing immersive virtual 3D navigation.

KEYWORDS: 3D urban modelling, mobile devices, mobile navigation, real-time visualization, virtual environments.

INDEX TERMS: I.3.4 [Graphics Utilities]: Application packages, virtual device interfaces, I.3.5 [Computational Geometry and Object Modeling]: Modeling packages, I.3.7 [Three-Dimensional Graphics and Realism]: Virtual reality I.3.8 [Applications]

1 INTRODUCTION

There is a growing need today for computer-based, photorealistic visualizations of 3D urban environments in many areas including environmental planning, engineering, telecommunications, architecture, gaming, 3D city information systems and even homeland security. Therefore, the procedural modelling of virtual cities in 3D is a topic not only computer graphics research but also other fields such as GIS or photogrammetry have focused on for a number of years. Different approaches have been favored with the two main ones being a fully-automatic or a semi-automatic one. It has been accepted that fully-automatic approaches are categorized according to the input information used.

To this day, the eldest and still an extremely popular data source for automatic virtual city generation, are aerial images [1].

Disadvantages of aerial images include the fact that analyzing information of this low level in dense urban contexts can be complex but it has been accepted that they can be a very rich source of various kinds of information. Automatic city generation has also been achieved via the use of laser scanning data [2]. This data, unlike aerial images, has the ability to represent the geometry of the surface directly. It has to be noted that the cost involved in acquiring this data is far greater than the one associated with aerial images, often almost prohibitive. However, technological advances since this type of data was first introduced have improved significantly with the cost decreasing progressively. The last method for automatic urban content creation, and one that has been increasingly popular of late, is the use of 2D ground maps/plans [3] from Geographical Information Systems (GIS). These can be very cost-effective in acquisition (often readily available) and are often used not just on their own but in conjunction with one of the previous data sources resulting in a more hybrid approach that can yield more accurate results.

It is commonly accepted that the ideal 3D urban modelling system would be a totally automatic one. Still, there are a number of obvious advantages to semi-automatic or user-assisted approaches where essentially an automatic system employs some limited operator input for guidance in a range of the tasks involved in order to offer better control on the resulting scene. Furthermore, the user-assisted (or interactive) tools can also be utilized as a more intelligent editing application for the preliminary model.

For example, a popular approach is to offer suggestions to an automatic urban modelling system which then in turn concludes the modelling task at hand on its own. This can lead to a more efficient method for modeling complex structures. A more conventional approach to interactive modeling is to provide the user with a set of generic models that are then adjusted (using usually image data) altering the model and the viewing parameters [4]. In this approach, the system provides geometric computations but the drawback is that substantial time and effort are required from the user.

Some other more recent approaches include the combination of user input with automatic processing introduced at various points with a different degree of control over the result. Offering an approximate building location to extract a building is one of the approaches suggested [5] that can lead to results although there is the important disadvantage of producing a final model very reliant on the automatic analysis. Other semi-automatic urban modelling tools have been described, some including methods for duplicating and/or cloning model meshes that are similar to others [6].

Yet another approach suggests an automatic system constructing topological relations amongst 3D roof points collected by a user for each roof. This method [7] results in a tool that can tackle easily several types of complex roofs. An additional notable interactive urban modelling approach involves the system handling complex building structures by means of constructive

¹email: c.gatzidis@soi.city.ac.uk

²email: vesna@soi.city.ac.uk

³email: f.liarokapis@coventry.ac.uk

⁴email: stuart.e.baker@alcatel-lucent.co.uk

solid geometry [8]. This system also uses an image correlation method to fit a primitive to the image. It should be noted however that this last method can be very costly on processing power when modelling more complex urban sites.

A particularly crucial area where 3D urban models are also in need today, apart from the ones above, is for urban navigation using mobile devices. This can range to a number of sub-applications including car and pedestrian navigation systems, location-based services and others. It appears that while the methods for automatic and semi-automatic creation of 3D urban models today do cover a lot of ground both in terms of concepts and also in delivery of content, they fail to take into account the considerable limitations and challenges mobile devices have to offer. Constraints in terms of graphics memory, processing power and memory cards render the output of most of these systems unusable on devices such as PDAs or smartphones because of excessive detail. At the same time, care needs to be taken in presenting this content on mobile devices since, for example, typical PDAs are only equipped with screens no larger than a 480x640 resolution display, thus presenting a further challenge.

We propose both a modelling solution (*Virtual City Maker*) that can deliver urban content efficiently to mobile devices as well as a virtual navigation environment (*Virtual Navigator*) on a PDA to accompany it so that that the resulting meshes can be put to the test. It should be noted that while the content is specifically optimised for mobile use in mind, for example a varying low-polygon count and adjustable texture image sizes, the results need to be of the highest quality possible rivaling virtual city models seen on desktop machines but also accurately placed in space to accommodate real-world navigational needs.

2 VIRTUAL CITY MAKER SOLUTION

In this section the overall workflow for the Virtual City Maker modelling tool will be described ranging from a description of the framework of the system to an overview of its *two modes*; the *automatic* and *semi-automatic* one with emphasis on some of the concepts, algorithms and interfaces behind them as well as some resulting 3D urban models.

Regarding the input data used (a very important part of the process which affects the results in many ways) the authors of this paper have selected a combination of aerial photographs and also 2D ground maps, at least for the outline of the buildings modelled. This more hybrid approach combines the strengths of two different techniques for a more efficient result. Aerial images ensure accurate results and the bypassing (to a certain extent) of a generalization process while ground maps provide positional and geographic information. Our work is not only focused on delivering content for mobile devices but also on navigation. Therefore, geo-referenced results and accuracy are key issues with errors in geometry placement only afforded to be marginal making this hybrid data use approach even more justified.

Having placed building footprint outlines in a precise manner, these outlines have been extruded to real world heights using, if needed, GIS vector data in the form of the .shp file format. This particular file format operates as a combination of CAD data with added metadata attributes. Some of these attributes include three different heights for each building (from which our system derives the average value, again when the height cannot be derived from shadow detection) plus additional useful data such as an individual ID for each building, its age, shape, type and others. To conclude the 3D urban creation process terrestrial photographs are also essential to construct a photorealistic 3D virtual city model. These terrestrial photographs are captured manually from

different viewpoints with the use of a conventional digital camera photographic technique.

Figure 1 illustrates in brief the modelling technique employed by the Virtual City Maker solution divided into several tasks. These tasks or methods are tackled by a unified system with one user-interface (*automatic mode*). They can also fall into different categories of work which means that they can also be resolved by individual modelling plug-ins (*semi-automatic mode*) that have been designed to also work as stand-alone applications in their own right.

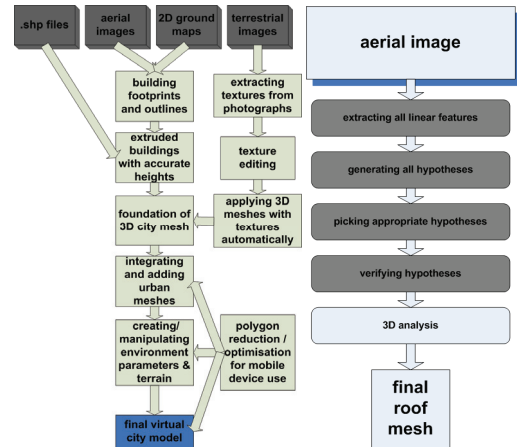


Figure 1: (a) Framework diagram of Virtual City Maker solution (b) Roof (and building) detection process block diagram

The key method in the *automated reconstruction mode* is the detection of building outlines and roofs from aerial images and/or 2D ground maps. The second other important step is the extrusion to real-world heights and that is handled, when needed, by a parser developed to read the GIS .shp files. Again, Figure 1 illustrates the main processes involved in the building and roof detection system.

The philosophy behind this has been to make only those decisions in the 3D analysis stage of the detection process that can be made in as much assurance at each level illustrated as possible. Therefore the hypothesis generation process creates hypotheses that may have somewhat weak evidence, which is where information from additional input data comes into use. Following that, a selection process then picks these based on more global evidence, filtering out the unwanted ones. The hypothesis verification process uses the most of this global information and can therefore make better decisions, resulting in a more accurate output. To begin with, initial hypotheses are generated for 2D projections of building roofs. The reason behind that is that it is generally accepted that roofs are usually the most distinct features in aerial images from a nadir point of view (such as the ones our system uses) and that focusing on the roofs helps reduce the number of hypotheses to be examined later on [11]. The generation of all roof hypotheses is gathered by the restriction of the building shapes to rectangular form or similar compositions and also of roofs to be flat. Roofs of buildings of this shape should project into a parallelogram-like shape.

The *interactive mode or user-assisted mode* of Virtual City Maker has been developed to take advantage of the fact that operator input can not only enhance the automatically-produced results and improve them but also diversify the appeal of the system itself. Often there are cases where users require urban content that is not derived from real-world GIS coordinates but can be fictitious (mobile gaming applications for example). The

semi-automatic mode of the application presented here can also cater for that too.

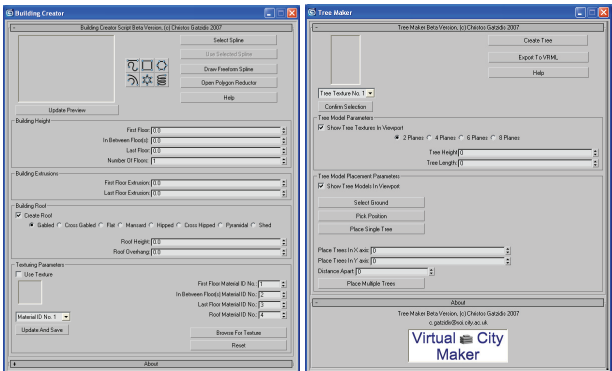


Figure 2: Some Virtual City Maker plug-in interfaces

This is achieved by dividing the modeling work to a number of plug-ins all created by scripting algorithms selected to efficiently tackle each issue at hand (some of the interfaces of the plug-ins are shown in Figure 2). These plug-ins at the moment include: the *Building Creator* script (results shown in Figure 3), which also includes a comprehensive building facade texture library, the *City Shader* script (results shown in Figure 4, designed to create different types of rendering modes that can be ported to our mobile engine, ranging from technical illustration to cartoon-shaded ones), the *Polygon Reductor* script (designed to optimize building structures while retaining their geographical shape characteristics, essential for navigation accuracy), the *Terrain Generator* script, the *Environment Maker* script (designed to enrich the urban scene with user-adjusted parameters such as lighting via manipulation of a sun-system orientation and others) and the *Tree Maker* script (designed to enhance urban scenes with foliage of adjustable and varying type).

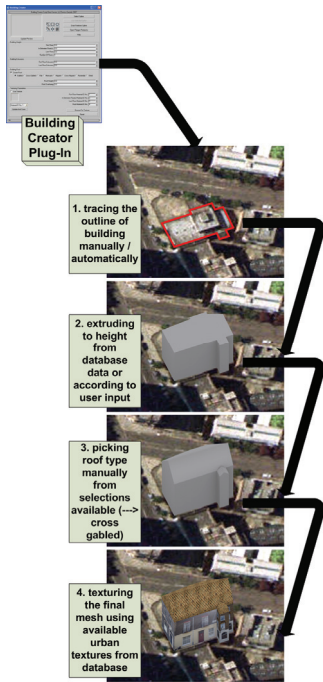


Figure 3: Step-by-step operation of Building Creator script

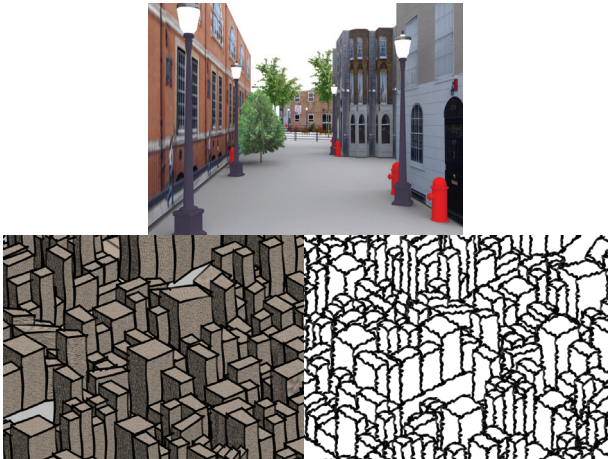


Figure 4: (a) Photorealistic, (b) Cross-hatching cartoon-shaded and (c) Distorted pen line renders using the City Shader script

3 VIRTUAL NAVIGATOR SOLUTION

The major component of navigation is travel and usually involves unconscious cognition whereas wayfinding is the cognitive process of defining a path via an environment using and acquiring spatial knowledge based on both natural and virtual cues [9]. Although a number of mobile navigation systems have been designed, they do not seem to have the appeal that was expected (the GPS in-car navigation systems aside). However, even these, have selected as the main visualisation environment to be the digital map which provides limited capabilities and they have now started to move into a rough 3D mesh visualisation which does not include photorealism. A brief overview of the most significant mobile VR applications has been recently documented [10]. Part of the problem with them lies on the provision of meaningful spatial information in a realistic manner, so that inexperienced users do not have to put a great amount of cognitive effort to ‘decode’ the retrieved information (i.e. 3D map or wayfinding operations). Another significant concern relates to the lack of user-friendly graphical interfaces that will allow for the customisation of the information provided as well as the provision of different navigation tools in real-time.

To address these issues, a VR pedestrian mobile environment called Virtual Navigator has been implemented. Previous user-studies [10] indicated that the use of photorealism in 3D urban maps used for pedestrian navigation is helpful and more attractive than the standard 2D digital map representation. However, a few users reported some problems with the interaction techniques used, mainly because of the limited functionality the interface provided. Another interesting point relates to the provision of choice to the user to accommodate sudden, external factors that may allow them to detour from a default path. In addition, it was positively suggested that the route line should be more distinct, minimising the probability of missing it while moving. To overcome these issues the graphical user interface (Figure 5) for user-centred navigation and wayfinding was re-designed from scratch and is now divided into four categories including *file*, *routes*, *directions* and *tools*. The ‘file’ category contains the necessary functionality to open and close geo-referenced spatial maps represented in 3D. The virtual maps are currently stored in the device by using the wireless connectivity (GPRS or WiFi) so the 3D maps can be downloaded from a web-server. This allows Virtual Navigator to meet one of the most significant requirements of modern navigation systems which is to be operational anywhere and anytime. The ‘exit’ option permits to

exit Virtual Navigator without the need of ‘killing’ the process from the memory control panel of the PDA. Next, the ‘routes’ category allows mobile users to select the type of representation they require for navigation and wayfinding operations. The effectiveness of the wayfinding depends on the number and quality of wayfinding cues provided to pedestrians [9]. Earlier user studies [10] indicated that users prefer different types of wayfinding aids like lines and arrows. In our work, this includes ‘arrows’, ‘lines’, ‘hotspots’ and ‘guides’. Arrows, lines and hotspots have been used individually in previous mobile prototypes [10] but not the guides. The purpose of using different categories of routes is to provide a meaningful aid that has a clear start and end, assisting pedestrians using Virtual Navigator, to find their way and not to get lost (size and colours of the arrows and lines can be customised).

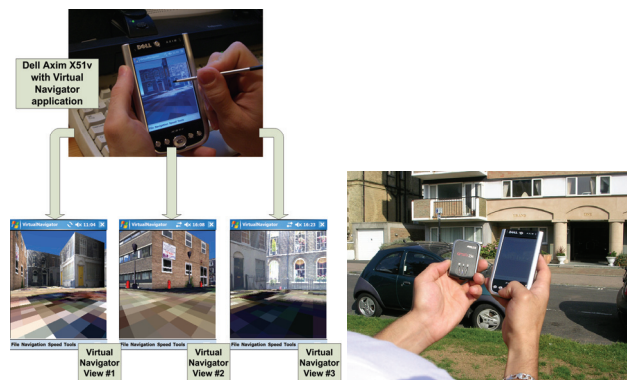


Figure 5: (a) manual mode (b) automatic mode

One of the points that were mentioned in previous user studies [10] is that the addition of recognisable landmarks would provide a clearer cognitive link between the VR environment and the real world scene. To address this effectively, first the 3D map was modelled using more detail such as trees and lamps which are considered as landmarks from a large number of users. Besides that, a number of hotspots that contain different types of functionality were added to the virtual 3D map making it interactive. In the simplest case, these include hyperlinks that link the 3D map with relevant web pages about the environment, but they can also provide links to other multimedia information such as digital pictures, audio and other 3D navigational information.

In our work ‘guides’ are virtual representation of humans, also known as avatars. Their main objective within Virtual Navigator is to guide pedestrians towards the requested destination. Currently, we have designed only one avatar, with male characteristics, which is activated as soon as the user clicks on him. Subsequently, the avatar starts walking with a slow pace, guiding this way the user, from one of the entrances of City University, London (St John Street) to the main entrance (Northampton Square).

The ‘directions’ category refers to the type of audio-visual information that can be provided to the pedestrians. The simultaneous presentation of audio-visual information meets one of the requirements (simultaneous activities that operate concurrently) of design issues in mobile interfaces. In the simplest case-scenario, textual directions provide information on how to navigate from one position to another. Additionally, textual annotations can be used effectively for presenting information about a place or a building (i.e. ‘...this is the main entrance of City University campus...’). Similarly, auditory directions perform the same action using pre-recorded way files, but are

prone to external noise. Finally, the ‘tools’ category, allows users to change some of the graphical and navigational properties of Virtual Navigator. These include the speed of navigation, the interaction type as well as the lighting of the scene.

To conclude, future work includes: a) the introduction of a plug-in, which via access to a library of different meshes can interactively enrich the virtual city scene with a number of objects (such as benches, pavements, street lamps etc.) further enhancing the scene considerably with user-assisted input b) the introduction of a new plug-in that can intelligently distinguish between rural areas and more complex urban areas etc. offering the user with a more “in-between” stage c) the introduction of a new plug-in that can semi-automatically model landmarks according to their type (statues, cathedrals etc.) d) integrating semantic details to buildings so that interior details can be modelled too e) conduct experiments with the existing results to gain insight on the issues of efficient modelling for mobile navigation and assess some of the usability issues of the interface f) implement a database with a management system to hold and maintain all the necessary geo-referenced spatial data into a secure server, serving the demands of current user-needs. Our server should feed the client, through an optimised network, with minimum latency in real-time.

ACKNOWLEDGEMENTS

Work presented in this paper was funded by ALCATEL Lucent Telecom UK Limited, the employees of which are also thanked for their input and insight.

REFERENCES

- [1] C. Jaynes, E. Riseman, and A. Hanson. Recognition and reconstruction of buildings from multiple aerial images. *Computer Vision and Image Understanding*, (90):1:68-98, 2003.
- [2] A. Wehr and U. Lohr. Airborne LASER Scanning - an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2):68-82, 1999.
- [3] G. Vosselman and I. Suveg. Map based building reconstruction from laser data and images. *Proceedings of Third International Workshop On Automatic Extraction Of Man-Made Objects From Aerial And Space Images*, (Ascona, Switzerland), pages 231-239, 2001.
- [4] T. Strat, L. Quam, J. Mundy, R. Welty, W. Bremner, M. Horwedel, D. Hackett, and A. Hoogs. The RADIUS Common Development Environment. *Proceedings of the DARPA Image Understanding Workshop*, (San Diego, CA), pages 215-226, 1992.
- [5] J. Li, R. Nevatia, and S. Noronha. User Assisted Modeling of Buildings from Aerial Images. *Proceedings of IEEE Conference On Computer Vision And Pattern Recognition*, (Ft. Collins, CO, USA), pages II:274-279, 1999.
- [6] Y. Hsieh. SiteCity: A Semi-Automated Site Modeling System. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, (San Francisco, CA, USA), pages 499-506, 1996.
- [7] A. Gruen and H. Dan. A Topology Builder for Automated Building Model Generation. *Proceedings of Automatic Extraction of Man Made Objects from Aerial and Space Images (II)*, (Birkhauser, Basel), pages 149-160, 1997.
- [8] E. Gülch, H. Müller, and T. Labe. Integration of Automatic Processes Into Semi-Automatic Building Extraction. *Proceedings of ISPRS Conference Automatic Extraction Of GIS Objects From Digital Imagery*, (Munich, Germany), 1999.
- [9] D. Downman, E. Kruijff, J. LaViola, and I. Poupyrev. *3D user interfaces: theory and practice*. Addison Wesley, pages 183-254, 2005.
- [10] C. Gatzidis, F. Liarokapis, and V. Brujic-Okretic. Automatic modelling, generation and visualisation of realistic 3D virtual cities for mobile navigation. *Proceedings of the 9th International Conference on Virtual Reality*, (Laval, France), pages 225-234, 2007.
- [11] C. Lin and R. Nevatia. Building Detection and Description from a Single Intensity Image. *Computer Vision And Image Understanding Journal*, 72(2):101-121, 1998.