

URBANPLANAR: BIM MOBILE VISUALISATION IN URBAN ENVIRONMENTS WITH OCCLUSION-AWARE AUGMENTED REALITY

Ludovico Carozza¹, Enrique Valero², Frédéric Bosché³,
George Banfill⁴, Rufus Mall⁵, and Minh Nguyen⁶

Abstract: Effective visualisation is critical to a wide range of activities conducted in the Architecture, Engineering and Construction (AEC) industry. This paper presents an *occlusion-aware Augmented Reality (AR)* solution that is geared towards the visualisation of new developments – in the form of 3D BIM models – in urban environments, to enhance communication around proposed developments. The solution automatically fetches BIM models from project repositories and enables a user to visualise them within their urban environment with a mobile tablet. The localisation of the tablet is achieved using a visual-inertial approach that uses only the sensors embedded in the tablet. Furthermore, the solution uniquely includes a process for handling occlusions from the real world on the augmented content. Results are presented that demonstrate the full pipeline, including some visualisation results obtained in real urban environments.

Keywords: BIM, Visualisation, Visual-Inertial, Urban Environment

1 INTRODUCTION

Effective visualisation is critical to a wide range of activities conducted in the Architecture, Engineering and Construction (AEC) industry, including for: a design team to communicate designs to clients and the public; a contractor to effectively explain a proposed construction programme to the client and sub-contractors; a repair and maintenance team to understand MEP circuits and their locations inside walls or under ceilings (Carozza et al. 2014a; Johansson et al. 2015; Côté et al 2013).

While construction data has been traditionally visualised through the production of 2D drawings, the increased availability of 3D (semantically-enriched) models resulting from the development of Building Information Modelling (BIM) is not just simplifying the production of 2D drawings, but enabling the use of new forms of visualisation techniques, such as Virtual Reality or Mixed Reality (Johansson et al. 2015; Mallia-Parfitt and Whyte 2014; Wang et al. 2013).

Virtual Reality (VR), whether experienced in a CAVE (Cruz-Neira et al. 1993; Mallia-Parfitt and Whyte 2014) or with immersive Head Mounted Displays (HMDs) (Johansson

¹ Research Fellow, School of Energy, Geoscience, Infrastructure and Society, CyberBuild Lab., Heriot-Watt University, Edinburgh, UK, l.carozza@hw.ac.uk

² Research Associate, School of Energy, Geoscience, Infrastructure and Society, CyberBuild Lab., Heriot-Watt University, Edinburgh, UK, e.valero@hw.ac.uk

³ Associate Professor, School of Energy, Geoscience, Infrastructure and Society, Leader of the CyberBuild Lab., Heriot-Watt University, Edinburgh, UK, f.n.bosche@hw.ac.uk

⁴ Technical Director, Linknode Ltd. Glasgow, UK, george.banfill@linknode.co.uk

⁵ Senior Software Engineer, Linknode Ltd. Glasgow, UK, rufus.mall@linknode.co.uk

⁶ BIM and GIS Engineer, Linknode Ltd., Glasgow, UK, minh.nguyen@linknode.co.uk

et al. 2015), enhances the level of immersion experienced by the user and enables the (collaborative) exploration of virtual environments at scale. While this supports more effective communication between stakeholders (who often do not all possess technical construction expertise), it also requires that entire environments be created digitally even though they already exist in reality. Furthermore, these technologies are not suited for use on site, where in-situ visualisation may in some cases be more meaningful. Mixed Reality (MR) solutions offer a solution to this.

Mixed Reality has actually been considered since the 1990's for its potential use in the AEC sector. For example, Feiner et al (1993) proposed KARMA (Knowledge-based Augmented Reality for Maintenance Assistance), a system to assist maintenance and repair tasks. Webster et al. (1996) reported experiments showing the location of columns behind finished walls. And, Szalavári et al. (1998) conceived a new system based on see-through HMDs, to immerse several users in an augmented world that makes use of markers to position augmenting objects. Ever since, computer-based visualisation has broadened to reach the whole lifecycle of buildings, from collaborative works during the design stage to modelling of design during the construction stage as well as subsequent maintenance operations (Bouchlaghem 2005). Portable and even wearable solutions have been developed aiming to increase the freedom of movement of users around virtual or augmented environments, improving users' immersive experience (Behzadan et al. 2015). Most notably is the Microsoft HoloLens see-through HMD with real-time indoor localisation tracking, which delivers Augmented Reality (AR) without specific markers (the room 3D environment acts as the marker). But, important limitation of technologies like the HoloLens is that they are meant to work in small room environments and may not be usable at larger outdoor scales (particularly because of the limitation in the range at which they can capture depth). Yet, the use of BIM for visualisation on site is gaining importance, including for urban planning purposes. Indeed, design BIM models could be used for in-field real-time visualisation, to better engage and improve the understanding of projects by stakeholders such as the public or city councillors and civil servants who make decisions on construction permits. Therefore, alternative solutions are required for AR outdoors.

In fact, with the advent of mobile computing, context-aware and ubiquitous visualization has become possible, with AR emerging as effective technology for simulation and monitoring of construction processes (Wang et al. 2013). Bae et al. (2014) present the HD4AR framework. In this system, infrastructure-less user localization is performed from single photographs, uploaded to a cloud-based service to query on-site relevant information. However, the HD4AR system is intended for on-site 3D authoring and annotation rather than for real-time continuous localization and augmentation. In contrast, Carozza et al. (2014b) proposed a visual-inertial sensor fusion framework to locate and track users navigating around augmented environments in real time.

In this paper, we present the pipeline of a novel real-time visualisation tool in which BIM models augment the scene in a live video recorded by a tablet computer. Similarly, to (Carozza 2014b), in the proposed system the augmentation is consistent with the static three-dimensional structure of the scene, so that occlusions (in particular, the real scene occluding the BIM) are correctly rendered. This functionality, which can be referred to as occlusion-aware Augmented Reality, is crucial to deliver the correct representation of the augmented scene to users. In contrast to (Carozza 2014b), we however deliver a system that runs in real time on consumer mobile devices (tablets).

2 OVERVIEW OF THE PROPOSED SYSTEM

In this paper, we present a full system that pulls BIM models from authoring environments, prepares them for mobile visualisation, and enables the user to view the models within their expected urban environment using nothing more than standard off-the-shelf mobile technology (i.e. consumer tablets). The user just needs to raise the mobile device, and then has freedom to move around to visualise the urban-inserted 3D design from locations of their choice. A distinct aspect of our system is that it takes account of occlusions of the existing urban environment (in particular other buildings) on the augmented visualisations.

The 6-DOF (three-dimensional position and orientation) tracking of the mobile device is conducted using a visual-inertial approach that relies solely on the camera and inertial measurement unit (IMU) sensors, now systematically embedded in these devices.

The overall system contains both offline and online processes that are summarised below and detailed in the following sections:

- Various **offline** processes are conducted to:
 - *Learn the visual and structural state of the existing urban environment (Section 3.1) with the generation of a geo-referenced 3D map of visual features and a geo-referenced model of the existing environment that is to be used as occlusion model.*
 - *Access the BIM design to be visualised and prepare it for efficient visualisation on mobile devices (Section 3.2).*
- During **online** operations (i.e. when the user holds the mobile device on site), a real-time visual-inertial algorithm finds and tracks the 6-DOF pose of the mobile device within the learned environment. This information is used to compute the view of the BIM model to be rendered onto the camera view as well as the occlusion mask to ensure that the view of the BIM model takes into account occlusions of the existing urban environment on it.

3 OFFLINE PROCESS

3.1 Learning the Existing Urban Environment

This process is summarised in Figure 1 and illustrated in Figure 2. This process requires that the urban environment be visited in advance to take a set of pictures from different locations (or a movie with travelling motion) covering the whole scene (see Figure 2a).

Key points with BRISK feature descriptors are automatically extracted in all images, and a Structure-from-Motion (SfM) algorithm (Snavely et al. 2008) is used to recover the relative poses between the different image cameras and the 3D coordinates of the key points, delivering a *sparse 3D point cloud* (see Figure 2b). The strength of each point in this point cloud is then evaluated and a *3D map of visual descriptors* generated following a process similar to the one described in (Carozza et al 2014b).

A multi-view stereo algorithm (Goesle 2007) then generates a *dense point cloud* (Figure 2c) that is subsequently used to generate a mesh by means of a surface reconstruction method (Fuhrmann 2014) robust to noisy input dataset (Figure 2d). This mesh is to be used during online processes, and its complexity (many small triangles that need to be processed) and approximations (straight lines may appear as broken lines; and some holes may exist within the mesh) could impact online processing speed and the

quality of occlusion calculations, respectively. Therefore, based on the common observation that the built environment is typically made up of planar surfaces, a simpler and more adequate mesh is instead generated by means of an automatic plane extraction algorithm applied to the dense point cloud (Figure 2e).

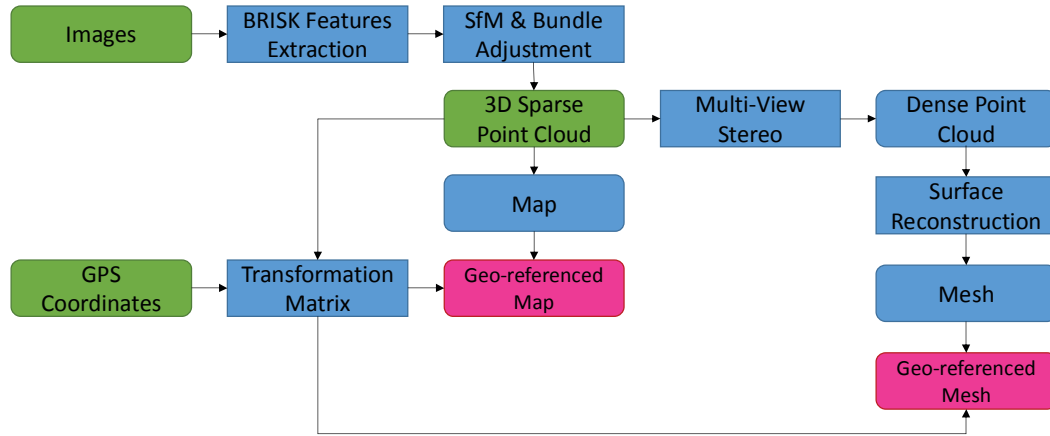


Figure 1: Outline of the offline process

One limitation of uncalibrated SfM models is the lack of absolute reference coordinate system as well as the unknown scale of the obtained models. Since BIM models are usually geo-referenced and in a metric scale, we also geo-reference (and thereby scale) the feature maps and occlusion meshes (Section 4). Geo-referencing offers an additional advantage that is also described later in Section 4.

3.2 BIM Model Retrieval and Preparation

The proposed system is designed to readily integrate within BIM processes. To this purpose, a web service has been created to allow users to manage their own data directly from their existing authoring pipelines. A plugin for Autodesk Revit has been developed to directly submit BIM models to the UrbanPlanAR cloud service. Once the models have been uploaded to our bespoke cloud service, they are processed using the open source “xBimToolkit”. This allows us to use the “xBim scripting language” to extract the external elements of the BIM model – disregarding the majority of the internal surfaces that would not be visible from street level. The resultant output from this stage - another IFC model - is then passed through the remainder of our pipeline, which includes a conversion to the *.obj* format allowing it to be imported into our rendering framework’s content pipeline. Here the model has final standard rendering optimisations applied to optimise for our targeted mobile devices.

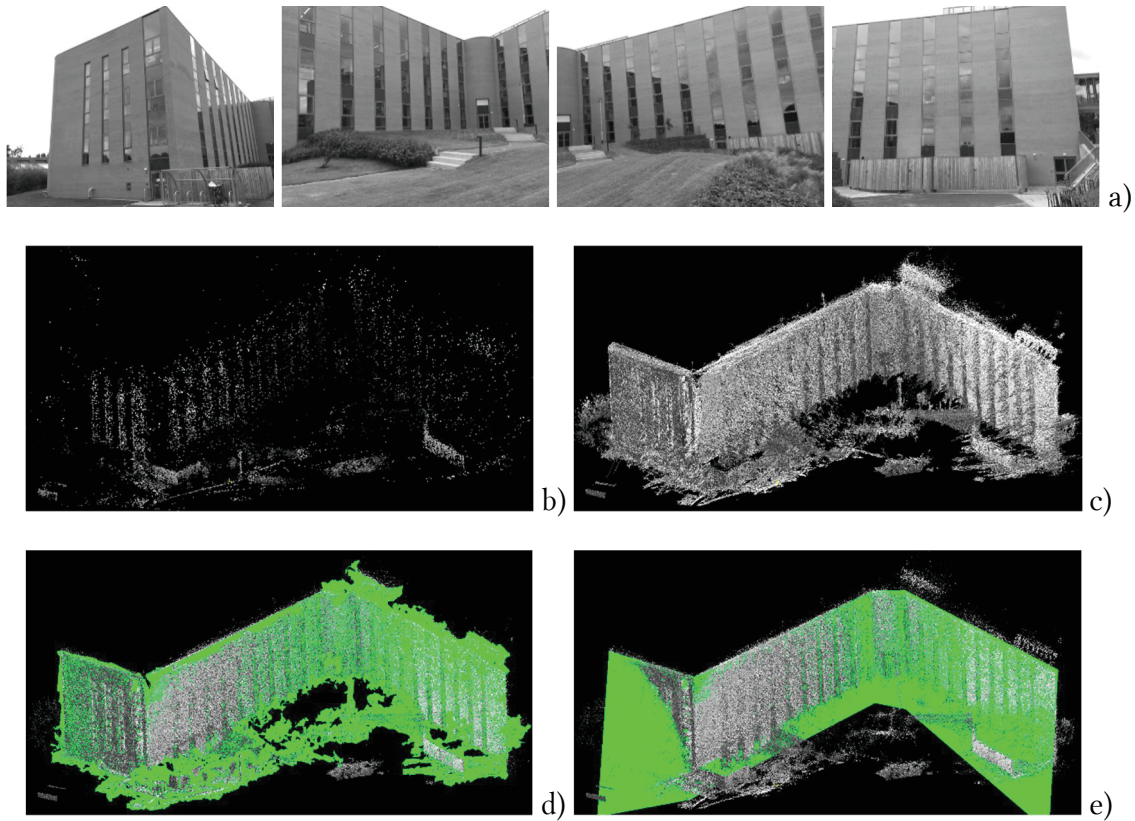


Figure 2: Different stages of the offline reconstruction. a) Picture/video frames of the environment b) Sparse point cloud. c) Dense point cloud. d) Occlusion mesh e) Occlusion mesh obtained from plane extraction

4 ONLINE PROCESS

During online operations on site, the real scene is acquired using the camera embedded in the mobile device and augmented in real time with the BIM model of the planned development. A visual-inertial 6DOF localisation algorithm tracks the camera pose as the user moves around (Section 4.1). This enables the accurate positioning and overlay of the augmenting 'BIM' model, including the accurate computation of its occlusion by the existing environment (Section 4.2).

4.1 Camera Localisation

A hybrid visual-inertial tracking approach has been developed to continuously determine the position and orientation of the camera in the geo-referenced system. The main process consists in extracting visual features (BRISK) from the camera feed and *match* them with the visual features stored in the feature map (see Figure 3a) to calculate the camera pose. To address jitter effects resulting from matching approximations and errors, this process is smoothed with an Extended Kalman Filter (EKF) that leverages additional inertial data from the mobile device's IMU. The IMU information is also used to maintain tracking when visual matching fails (e.g. during fast movements). This approach is an improved version of the algorithms described in (Carozza et al. 2014a) and (Carozza et al. 2014b). In particular, a new constant velocity differential model has been used for the EKF to take into account the higher noise of rotations measured from tablets' IMU. Feature tracking has also been optimized implementing efficient periodical re-matching

of features. For this, we keep track of a "good" subset of features with high repeatability and matching chances in the background. This improves stability and limits the accumulation of drift errors while preserving real-time performances.

4.2 Rendering with Occlusions

For each frame, the camera pose calculated by the camera localisation algorithm is used to render the view of the BIM model. Using the occlusion model (that is automatically aligned with the view during localisation), depth-buffering is employed to determine which parts of the BIM model should indeed be visible and rendered, and which parts are actually occluded by the real scene and should thus not be rendered (Figure 3b).

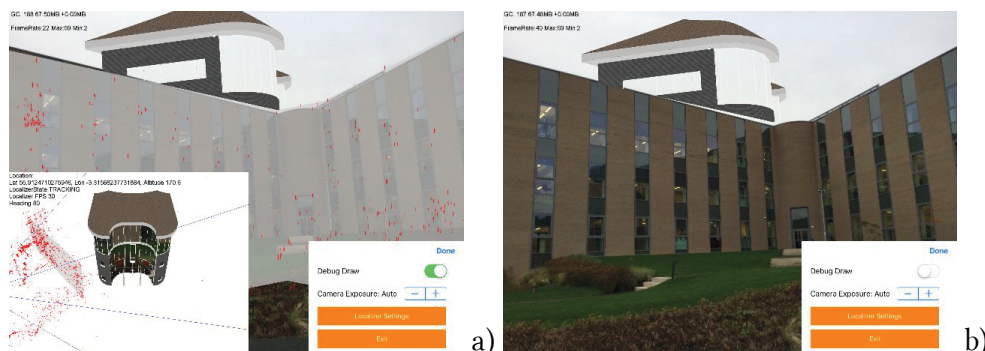


Figure 3: Online processes. a) Camera localisation. This view shows the map of visual features overlaid on the current view. b) Rendering of BIM model on top of the image with correct occlusion calculation.

5 RESULTS

The proposed system has been tested in several environments under different weather conditions with truly variable lighting conditions (between 300 and 20,000 lux). Amongst them, some trials were carried out around the Lyell Centre at Heriot-Watt University in Edinburgh (UK) and their results are presented here. The Lyell Centre is a L-shaped building with a length and width of ~30m and a height of 12m (see Figure 2).

To carry out the offline process, the researchers first visited the facilities acquiring a video of the site using an iPad Air 2 with a camera resolution of 640x480 pixels. Around 200 video frames were automatically selected for the offline stage. These pictures were processed to output a map of 3,350 3D visual features, finally reduced to 2,250 points, as explained in Section 3.1. Following a Multi-View Stereo process, a dense point cloud with 11 million points is also created that is subsequently converted into a mesh with 1.2 million faces. This mesh was finally converted into a simpler model, with just 4 patches fitting the main walls (see Figure 2). The BIM model was then exported from Autodesk Revit and processed using the pipeline mentioned in section 3.2.

During the online process (we use an iPad Air 2 running iOS 9.3; but our application also runs on Windows devices), the mobile application was launched and the tablet pointed towards the site. The application then tracked the 6DOF location of the tablet in real-time (approx. 30 fps) and augmented the view, as described in Section 4.

Figure 4 shows some augmented frames from two live videos related to two different test motions carried out starting from different locations. During the experiments, the user was moving around the building, at a distance of 20 m approximatively, for about 2

minutes. Videos of this and other experiments can be found at <https://web.sbe.hw.ac.uk/fbosche/projects-urbanplanar.html>.

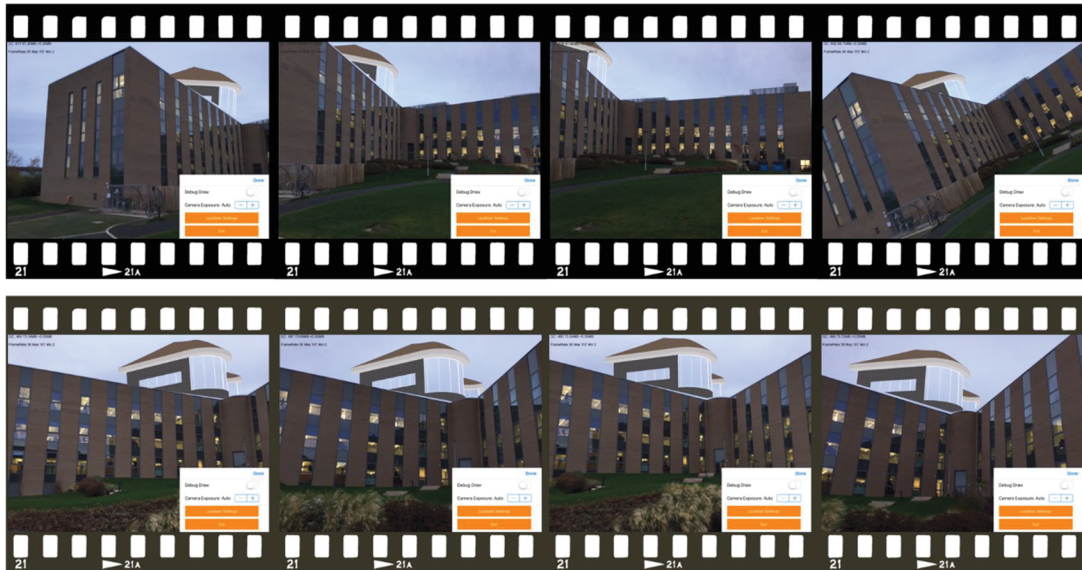


Figure 4: Frames from two online tests carried out in Lyell Centre.

6 CONCLUSIONS

This paper presented an end-to-end solution, called *UrbanPlanAR*, for BIM mobile visualisation in urban environment with occlusion-aware augmented reality. As opposed to other Virtual Reality solutions, *UrbanPlanAR* enables users to visualise forthcoming developments in the real environment, delivering a more realistic, empowering experience at a potentially lower cost (less modelling required). Furthermore, in contrast to the vast majority of current AR solutions, *UrbanPlanAR* includes the important capability of taking static occlusion of the existing urban environment into account when augmenting the view with the virtual content. Altogether, this solution will be of great use to various stakeholders of construction projects to enhance their understanding of the impact of new developments on the urban landscape. Future work will investigate alternative means to more effectively compute and geo-reference the offline data, as well as even more robust online tracking methods.

7 ACKNOWLEDGMENTS

This work was supported by Innovate UK [Grant number: 10-20-40].

8 REFERENCES

- Bae, H., Golparvar-Fard, M., White, J. (2014) Image-based localization and content authoring in structure-from-motion point cloud models for real-time field reporting applications. *J Comput Civil Eng. ASCE*, Reston. B4014008.
- Behzadan, A. H., Dong, S. and Kamat, V. R. (2015) Augmented reality visualization: A review of civil infrastructure system applications. *Advanced Engineering Informatics* 29 (2015), 252-267.

- Bouchlaghem, D., Shang, H., Whyte, J., and Ganah, A. (2004). Visualisation in architecture, engineering and construction (AEC), *Automation in Construction* 14 (2005), 287-295.
- Carozza, L., Tingdahl, D., Bosché, F. and van Gool, L. (2014a). Markerless Vision-Based Augmented Reality for Urban Planning, *Computer-Aided Civil and Infrastructure Engineering*, 29: 2-17.
- Carozza L., Bosché F., Abdel-Wahab M. (2014b). Robust 6-DOF immersive navigation using commodity hardware, Proceedings of the 20th *ACM Symposium on Virtual Reality Software and Technology (ACM VRST 2014)*, Edinburgh UK, pp. 19-22.
- Côté S., Barnard J., Snyder R., Gervais R. (2013). *Offline spatial panoramic video augmentation for visual communication in the AEC industry*. International Conference on Construction Applications of Virtual Reality (CONVR), London, UK.
- Carolina Cruz-Neira, Daniel J. Sandin and Thomas A. DeFanti (1993). *Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE*, SIGGRAPH, pp. 135-142.
- Dalton, B., and Mallia-Parfitt, M. (2013). Immersive Visualization of Building Information Models, Design Innovation Research Centre, Working Paper 6.
- Fuhrmann, S., and Goesele, M. (2014). Floating Scale Surface Reconstruction, *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2014)*, Vancouver, Canada, 2014.
- Goesele, M., Snavely, N., Curless, B., Hoppe, H., and Seitz, S. M. (2007). Multi-View Stereo for Community Photo Collections, *Proceedings of ICCV 2007*, Rio de Janeiro, Brasil, October 14-20, 2007.
- Johansson, M., Roupé, M., Bosch-Sijtsema, P. (2015) Real-time visualization of building information models (BIM), *Automation in Construction*, 54: 69-82.
- Mallia-Parfitt, M., and Whyte, J. (2014). Developing a mobile visualization environment for construction applications, Proceedings of the 2014 International Conference on Computing in Civil and Building Engineering, Orlando, Florida, USA, 23-25 June 2014.
- Snavely, N., Seitz, S.M., Szeliski, R. (2008) Modeling the World from Internet Photo Collections, *International Journal of Computer Vision* (2008) 2:80, 189-210.
- Szalavári, Z., Schmalstieg, D., Fuhrmann, A., and Gervautz, M. (1998). "Studierstube": An Environment for Collaboration in Augmented Reality, *Virtual Reality* (1998) 3:37-48.
- Van den Bergh M., Halatsch J., Kunze A., Bosché F., Van Gool L., Schmitt G. (2009). A novel camera-based system for collaborative interaction with multi-dimensional data models, *Proceedings of the 9th International Conference on Construction Applications of Virtual Reality - CONVR 2009*, November 5-6, 2009, Sydney, Australia.
- Wang, X., Mi Jeong, K., E.D. Love, P., Shih-Chung, K. (2013) Augmented Reality in built environment: Classification and implications for future research, *Automation in Construction*, 32:1-13,
- Webster, A., Feiner, S., MacIntyre, B., Massie, W., and Krueger, T. (1996), Augmented reality in architectural construction, inspection and renovation, *Proceedings of ASCE Third Congress on Computing in Civil Engineering*, Anaheim, California, USA, 1996.