Tracking Visual Engagement in Mixed Reality: A Framework for Analysing Interaction with Digital Models

Adam Tomkins[1], Sigrid Hehl-Lange[2], Eckart Lange[2]

[1]The University of Sheffield, Sheffield/UK · a.tomkins@sheffield.ac.uk
[2]The University of Sheffield, Sheffield/UK

Abstract: Due to technological advances, affordable mixed reality devices, including both virtualreality and augmented reality, are beginning to be used to support participatory planning to better inform stakeholders of design interventions. Stakeholder participation and perception studies often utilize both qualitative interviews and quantitative techniques to gauge how a model can effectively communicate design intentions. In this paper, we examine how mixed reality enables new methods of investigating how users interact with digital models, enhancing existing techniques.

We introduce a non-invasive perspective tracking system, to track how a participant explores a 3D model of a landscape intervention through the perspective of their mixed reality device. Our software provides a supplementary technique to better analyse user interaction patterns with digital models, allowing a data-driven approach to support participation studies and enable novel analytical approaches. We introduce the concept of naturally salient perspectives and show that tracking how a user frames a model during free exploration can provide new quantitative insights to support traditional qualitative participation methods.

Keywords: Stakeholder Participation, visualization, augmented reality, virtual reality

1 Introduction

In landscape architecture, visual representations form the primary means of communication between stakeholders, driven by advances across a variety of media, from the iconic before-and-after paintings in Repton’s red books, to computer-generated images and digital models.

Digital models have proven to be an adaptable medium to demonstrate landscape interventions, with the ability to combine the spatial features of traditional scale models, with temporal features such as the progression of new developments, seasonal variations and guided explorations (LANGE 2011). Building on this, mixed reality (MR) visualizations are particularly suited to interacting with augmented layers of contextual information (GHADIRIAN & BISHOP 2008) and enable a more informative model exploration experience.

As a medium to convey complex ideas, the ability of a digital model to efficiently communicate design intentions can be better understood by studying how participants interact with it. We suggest that while mixed reality may be used to help stakeholders to understand a proposed intervention, it may also enable researchers and designers to learn about how their models are consumed by their participants. Mixed reality encompasses both virtual reality, in which a participant is free to explore an entirely simulated environment, and augmented reality, in which a participant can explore a simulated environment which is situated within the real world.

Journal of Digital Landscape Architecture, 4-2019, pp. 160-167. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-663-5, ISSN 2367-4253, doi:10.14627/537663017.

Due to technological advances, affordable mixed reality devices are beginning to be used to support the participatory planning process and to better inform stakeholders of design interventions in on-site and off-site sessions (PORTMAN et al. 2015, WANG 2009). GOUDARZNIA et al. (2017) use a tablet-based mobile application to test the effectiveness of augmented reality in the public participation process. They demonstrate that, as part of an on-site presentation, participants feel comfortable with using augmented reality as a tool to explore future interventions.

Shelton (2003) demonstrates that using augmented reality can help participants learn dynamic spatial relationships. This is further reinforced by Soria and Roth (2018), who show that by using augmented reality to engage our innate spatial cues through locomotion, they can improve a participant's spatial cognition when asked to recall the specifics of a proposed landscape intervention. Mixed reality can combine both the intuitive interactions of real-world models, with fine-grained interactions possible with digital models, enabling a more in-depth analysis of participants interaction. Research into participation and perception often utilizes both qualitative interviews (SORIA & ROTH 2018, GOUDARZNIA et al. 2017) and quantitative techniques such as eye tracking (DUPONT et al. 2014). Eye tracking technology provides an accurate measure of the eye's saccadic movements toward naturally salient features across a visual scene, such as wind farms, or bodies of water. Traditional research using eye-tracking has been successful in elucidating how users actively perceive landscape in 2D. Yet real-time eye tracking remains intrusive and ill-suited for naturalistic exploration of 3D environments.

In this paper, we propose a mixed reality counterpart to the salient feature detection in eye tracking, wherein we track the naturally salient perspectives unveiled during the free exploration of a 3D model. Through this device-based perspective tracking, we propose instead to track how a participant naturally chooses to frame a scene by analysing the array of participant-generated visual perspectives created during visual interactions. This research seeks to demonstrate the potential of mixed reality as a tool in landscape architecture research and practice. By researching the opportunities afforded by the adoption of virtual and augmented reality technology, we seek to encourage novel stakeholder participation methods utilizing mixed reality, and facility the adoption of mixed reality in practice.

2 Methods and Implementation

2.1 Design Principles of the Perspective Tracking Framework

We designed the Perspective Tracking Framework with the aim of enhancing stakeholder participation. We aim to harness immersive technology as both a qualitative source of information and a tool to support further quantitative enquiry, suitable for an expanding array of landscape architecture use-cases. To support this aim we record a variety of data including:

- the total time spent interacting with the model, from each vantage point,
- the total time spent focused upon each point in the proposed model,
- the unique number of visits to each vantage point and focal point.

With these data, we can reconstruct a variety of interaction phenomena to document how stakeholders participate with planned design solutions.

The framework must also support interview questions to backup qualitative studies, and thus be able to process and display insights both during experiments, and supply feedback as soon as an experiment has finished. Previous studies have relied upon briefing participants to take screenshots of interesting areas in a proposed intervention (SORIA and ROTH 2018), this framework must complement these self-reporting approaches with an objective assessment of the most salient perspectives.

In order to be a widely applicable framework, it must also be able to run on the majority of platforms used in virtual and augmented reality landscape research. To support both on-site and off-site experiments, the framework must support desktop computers, dedicated headsets such as the Microsoft HoloLens and Oculus Rift, and mobile devices such as tablets and phones. Finally, the software must be generic and configurable, in order to be re-usable across experiments, as well as be easy to implement and expand for new use cases as the technology progresses. It will not be limited in scope to the use-case presented in this paper.

2.2 Design Elements of the Perspective Tracking Framework

Figure 1A illustrates the 5 main elements required for the Perspective Tracking Framework: *viewers, models, viewports, focal points and perspectives.* Primarily each experiment will require one or more *viewers* and a digital *model* with which to interact. Each viewer will require a mixed reality device which will provide the *viewport* through which to see the model. This could be a mobile device providing the camera as the viewport, a headset which provides the viewport on head mounted displays, or a virtual camera rendering a scene to a desktop computer.

From these elements, the viewport can determine the *focal point* of the viewport on the model, as defined by where the central point of the perspective meets the model. It is worth noting that 'perspective tracking' here only considers the perspective of the user's device as defined by the viewport. We do not track individual eye movements. By periodically recording the location and progression of the viewport and focal point through space, we determine the natural flow of *perspectives* that a viewer chose to follow as they visually consume a model. From this we may reconstruct and dissect the visualization experience.

2.3 Implementation of the Perspective Tracking Framework

We implement the framework in two stages. Firstly, we have the software required to track the visual interactions with a model in real time. Secondly, we have the software required to explore the data generated by the tracking and analyse the experimental results. Both are wrapped into a single package available for the Unity Game Engine. We chose Unity as it is the standard platform used to develop mixed reality applications, currently supporting a growing body of work within the landscape architecture literature (GOUDARZNIA et al. 2017, HAYNES et al. 2016, HAYNES et al. 2018, SORIA & ROTH 2018). We demonstrate this framework in use by creating a mobile application for exploring proposed developments in the Pearl River Delta (Figure 1B).

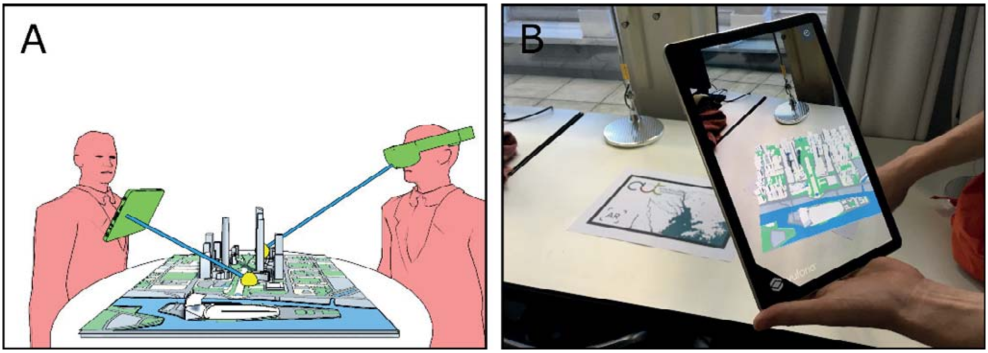

Fig. 1: An illustrative Augmented Reality example and digital model of the Perspective Tracking Framework. **A** A mock-up of the framework components. Viewers are shown in light red, with their AR devices in green. The perspective's focal point on the digital model of each device is shown in yellow. **B** An example application with an android tablet. The tablet screen shows the viewer's perspective of the model, which is overlaid onto the tracking marker seen on the table.

The relative positions of the viewport and the model can be derived from either headset tracking in virtual reality or environmental tracking in augmented reality. As the user moves through space and changes their perspective, we record the time spent in each location, as well as the perspective's focal point on the model. This allows us to aggregate spatial attention patterns as the participant interacts with the model. From this, we can reconstruct a participant's perceptual experience and collate visual attention coverage to extract perspective maps. Condensing continuous participant trajectories down to a discrete three-dimensional grid system of 'voxels' provides easily visualised and interpreted data. As such we divide the viewport space and the model space into two separate configurable grids, shown in Figure 2. The viewport grid (shown in yellow) is generally divided into larger voxels (~15 cm^3), while the model grid (shown in red) is often much smaller (~3 cm^3) enabling higher resolution to record the saliency of specific features.

2.4 Technical Requirements of the Perspective Tracking Framework

The software is designed to be non-intrusive and incorporated into the current experimental paradigms. It is a Unity Game Engine package, working alongside current mixed reality toolkits such as Vuforia, Google ARCore, Apple ARKit and Microsoft HoloLens for augmented reality, as well as SteamVR and Google Cardboard, supporting all major virtual reality headsets. The framework adds minimal computational overhead and can be run without issue on mid-range mobile phones. For model visualisation, we support common modelling formats including SketchUp, Wavefront OBJ, FBX, as well as 360° images.

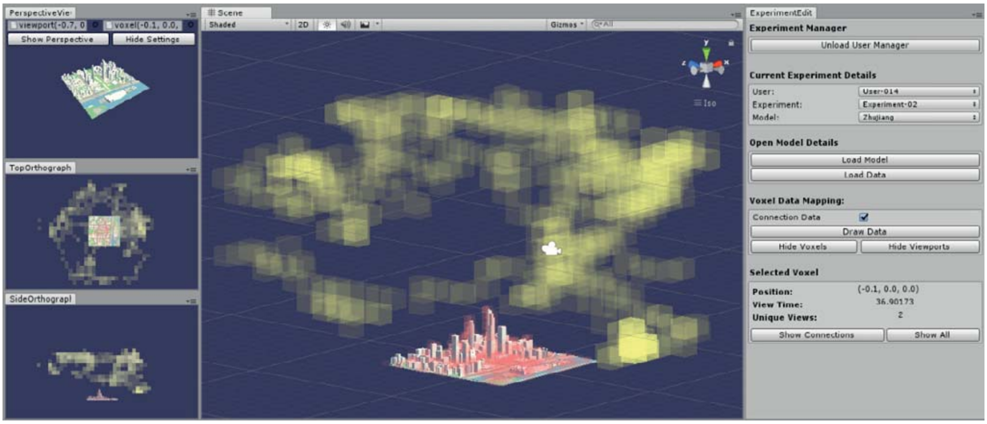

Fig. 2: The data analysis interface in the Perspective Tracking Framework. The main window shows the results of a single session of free model exploration. Yellow voxels indicate the viewport positions, red voxels indicate the focal points on the model. The left window has a perspective camera, recreating a specific perspective, from a selected pair of viewport and focal point. This is followed by two orthographic views. On the right is the editor extension controlling the data in the experiment. Increased opacity in voxels reflects increased number of connections between focal point voxels and viewports.

3 Results

In this paper we have described a theoretical framework, implemented as an editor extension and interface within the Unity Game Engine, as shown in Figure 2. We create a mobile application to explore proposed developments in the Pearl River Delta in augmented reality (Figure 1B) to demonstrate one possible use case of the framework. This provides a wealth of model interaction data which detail how a model is being visually consumed. We introduce the notion of 'naturally salient perspectives' to address a range of data-driven questions, such as:

- What design elements promote investigation from the multiple perspectives, and which viewpoints are exploited for the most varied views?
- Which areas are transiently frequented, and which captivate for longer periods?
- Which features prompt a closer inspection or encourage a distant impression?

3.1 Data Analysis in the Perspective Tracking Framework

Perception maps provide an intuitive and informative summary of model exploration**.** From the data requirements set out in section 2.1, we overlay data onto the model to visually represent the spatial distribution of perspectives around a model. Figure 2 shows the full spatial dataset recorded in a single trial, illustrating the exploratory patterns exhibited during model interaction. Figure 3 shows two examples of how we summarise specific visual interactions into perspective maps. A perspective map can come in several forms, representing either a spatial snapshot of a participant's attention originating from a single position, as shown in

Figure 3A, or alternatively displaying attention patterns with a central focal point on a design feature, as shown in Figure 3B.

Maps can display a variety of statistics gathered to answer different questions**.** The appearance of the viewport voxels can be configured to show: the total time spent viewing from that vantage point, the number of times the viewer returned to that vantage point and the number of different focal points associated with that vantage point. The focal point voxels can be configured to show their counterpart of the same data. Figure 3B shows how we combine a temporal feature, the total viewing time, and spatial features, the distribution of different perspectives centred on a feature, to provide experimental insight (section 5).

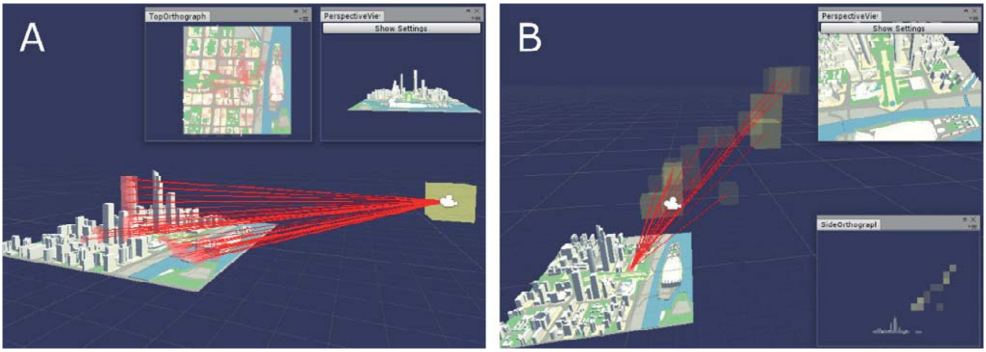

Fig. 3: Perspective maps display spatial and temporal behaviours. **A** Model exploration from a single viewport showing the most used vantage point to focus on the central towers. **B** A feature focused map showing all the viewports with perspectives centred on a central axis building. Voxel opacity encodes time spent viewing from each voxel. Inset is the user's perspective from a representative viewport-voxel pair, and an orthogonal view. Red lines show perspective-focal point connections.

Automated perspective reproductions can recall salient perspectives to support follow-up interviews**.** Inset into each map in Figure 3 is a reproduction of a representative salient perspective, as determined by its increased viewing times with respect to the whole dataset. A perspective displays what would have been seen on the mixed reality device in use. Figure 2 shows additional orthographic perspectives of the visual interaction distributions, which provide a 2-dimensional summary of the overall consumption patterns of the model, used to quickly identify areas of interest, such as the main axis.

These data can be used as part of a follow-up analysis, as either an analytical dataset or as a quantitative basis for guiding a qualitative study on perception and user interaction. We support real-time visualizations and offline analysis for data analysis across multiple users, experiments and models, enabling large cohort studies and robust statistical analysis.

4 Discussion

Primarily we consider how the adoption of this framework can enhance the role of visualizations in stakeholder participation practice, with respect to Arnstein's ladder of stakeholder participation (ARNSTEIN 1969). We argue, that through a more immersive medium, mixed reality can enable a more reciprocal participation process in both experimental studies and design consultations. For participants, mixed reality can serve as a tool to enhance spatial understanding and provide context to proposed interventions. Using this framework, participants can then feed back into the design decisions through two paths.

During experiments, this framework provides a mechanism to more constructively guide an interview process, by enabling the experimenter to recall a participants naturally salient perspectives to provide direct feedback through examination, reflection and input. This can enhance previous methods of depending on screenshots to note interesting features, thereby driving questions relating to well explored features (SORIA and ROTH 2018). Indirect feedback can be provided to the designer through a visual representation of how a model is consumed. For example, Figure 2 shows that there are 2 main modes of interaction for this model. The orthographic views show two rings of voxels at different elevations and radii.

The upper ring provides a bird's eye view of the model, however, more subtly there are areas at a lower elevation, which show vantage points where the user has stooped to get a 'skyline view' of the model. This is supported by Figure 3A which shows that the view across the river and up the main axis, showing both towers, forms a naturally salient perspective displayed in the inset perspective view. The inset top-down view shows the relative increase in focal points surrounding the main axis. This is reinforced in Figure 3B, wherein the side-view clearly shows increased time spent focusing closely on the area, specifically on the blue and green spaces where the main axis meets the river, as demonstrated by the reconstructed perspective from the most salient viewport. These insights provide feedback to influence consultation practices. For instance, areas which are shown to be overlooked could indicate a need for visual improvement on the part of the designer or merit a more detailed discussion in stakeholder meetings to ensure the proposed changes in overlooked areas are not subject to a failure of communication.

This paper has highlighted a single use case of the framework, which is typical of demonstrating large scale planning designs. However, this work is applicable to both research into the participation processes itself, and more generally as a tool for existing public participation processes. This framework can be further used in more exploratory methods such as simulated walks through proposed parks, or visual impact assessment in 360-degree panorama. Further applications of this work could include architectural exploration, environmental planning, evaluation of marketing materials, or collaborative digital design.

5 Conclusion

In this paper we propose that new methods of capturing user interaction in mixed reality provide a rich vein for further research topics and enable new research questions in the domains of participation, visual engagement, and visualisation. We detail the creation of an open framework for mixed reality experiments, which supports a large array of possible experimental scenarios. Finally, we develop a mobile application to monitor visual interactions with a proposed landscape intervention in the Pearl River Delta, demonstrating the prospect of data-driven perspective mapping during stakeholder participation studies.

This research is supported by the Newton Fund / Engineering and Physical Sciences Research Council (EPSRC) as part of the Sustainable Deltas award. It is a collaborative project with South China University of Technology, who provided the Sketchup Model used in our research, and Delft University of Technology (TU Delft).

References

ARNSTEIN, S. (1969), A Ladder Of Citizen Participation. Journal of the American Institute of Planners, 35 (4), 216-224.

DUPONT, L., OOMS, K., ANTROP, M. & VAN ETVELDE, V. (2017), Testing the validity of a saliency-based method for visual assessment of constructions in the landscape. Landscape and Urban Planning, 167, 325-338.

GHADIRIAN, P. & BISHOP, I. (2008), Integration of augmented reality and GIS: A new approach to realistic landscape visualisation. Landscape and Urban Planning, 86 (3-4), 226-232.

GOUDARZNIA, T., PIETSCH, M. & KRUG, M. (2017), Testing the Effectiveness of Augmented Reality in the Public Participation Process: A Case Study in the City of Bernburg. In BUHMANN, E., ERVIN, S., HEHL-LANGE, S. & PALMER, J. (Eds.), Journal of Digital Landscape Architecture, 2-2017, 244-251.

HAYNES, P., HEHL-LANGE, S. & LANGE, E. (2018), Mobile Augmented Reality for Flood Visualisation. Environmental Modelling & Software, 109, 380-389.

HAYNES, P. & LANGE, E. (2016), Mobile Augmented Reality for Flood Visualisation in Urban Riverside Landscapes. In BUHMANN, E., ERVIN, S., HEHL-LANGE, S. & PALMER, J. (Eds.), Journal of Digital Landscape Architecture, 1-2016, 254-262.

LANGE, E. (2011), 99 volumes later: We can visualise. Now what? Landscape and Urban Planning, 100 (4), 403-406.

PORTMAN, M. E., NATAPOV, A. & FISHER-GEWIRTZMAN, D. (2015), To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning. Computers, Environment and Urban Systems, 54, 376-384.

SHELTON, B. (2003), How augmented reality helps students learn dynamic spatial relationships. PhD thesis, Wageningen.

SORIA, C. & ROTH, M. (2018), Unreal Reality : An Empirical Investigation of Augmented Reality Effects on Spatial Cognition in Landscape Architecture. In BUHMANN, E., ERVIN, S., HEHL-LANGE, S. & PALMER, J. (Eds.), Journal of Digital Landscape Architecture, 3-2018, 150-162.

WANG, X. (2009), Augmented Reality in Architecture and Design: Potentials and Challenges for Application. International Journal of Architectural Computing, 7 (2), 309-326.