8

Conclusions

The techniques presented in this dissertation draw from research areas including mixed reality, 3D reconstruction and tabletop display interaction, to facilitate a novel way to communicate situational and navigational information. This dissertation makes a number of significant and original contributions in the area of remote communication.

The remainder of this chapter summarises the novel contributions of this dissertation. These contributions address the goals and research questions in the introduction to this thesis. This chapter then discusses the future work that will be undertaken to further improve the system presented in this dissertation. It will also discuss research opportunities made possible by the techniques and system presented in this dissertation. This chapter finishes with some concluding remarks.

8.1 Contributions

God-like interaction is an intuitive way for people in a command and control center-like environment to provide situational and navigational information to remote field workers. Inspired by popular culture representations of various gods, the techniques leverage the natural inclination of humans to use gestures and physical prop interaction to help express ideas and instructions.

8.1.1 Observations

The research presented in this dissertation shows that instructions and intentions are well conveyed using the god-like interaction techniques. However, distance estimation was poor: virtual objects are always rendered in front of real world objects regardless of distance as there is no depth information captured by the system. There is also evidence to suggest that a change in scale of HOG objects is predominantly incorrectly perceived as a change in distance.

Gesture-based collaborative navigation was shown to be significantly more efficient than audio only collaboration. While not significantly different to other forms of visual-based collaborative navigation, it does facilitate experimentation and exploration. There is also some evidence to suggest that without audio cues, gesture-based communication is faster than mouse-based communication for collaborative navigational tasks due to the extra ambient visual information.

8.1.2 Hardware

The HOG table is a novel tabletop display capable of 3D capture on a scale not previously researched. It facilitates a number of people working shoulder to shoulder in a collaborative manner. While the display facilitates visualisation of areas of real or virtual worlds, the 3D capture system affords natural hand gestures and tangible prop interaction to be performed with a geospatial reference. As the system does not attempt to interpret the interactions (but simply conveys them "as is") the system does not restrict users in the manner in which they can interact.

8.1.3 Software

To support the rendering of 3D reconstructed objects on an outdoor AR system, the VBR algorithm presented by Li et al. (2003a) was extended to work over an unreliable low-bandwidth network.

8.2 Future work

The system presented in this dissertation facilitated exploration of the practical uses and limitations of god-like interaction for remote communication. There are many encouraging results, but also some restrictive aspects that need addressing before it is can be considered beneficial. The remainder of this section discusses future improvements to the system as well as new research opportunities that exist because of the research undertaken in this dissertation.

8.2.1 Visual hull

The 3D reconstruction approach used in this dissertation restricts the number of cameras that can be used to one more than the number of texture units on the graphics card. Commonly graphics cards only have four texture units, restricting the number of cameras to five. With such a small number of view points HOG objects only roughly resemble the original shape of the object. Using a multi-pass approach (Li et al., 2003b) it is possible to use an arbitrary number of cameras.

Typically, using the standard VBR approach (Li et al., 2003a) a seam is visible where textures from different viewpoints meet. Using a shader language, such as Cg (Mark et al., 2003), the seams and other aliasing artefacts can be eliminated (Li et al., 2004).

The current system is able to perform 3D capture at seven FPS. This is a limitation of processing camera images on the CPU. These times could be substantially improved by offloading this processing onto the GPU as much of the processing would be performed in parallel. Another alternative is cameras with on-board field-programmable gate arrays (FPGAs). An FPGA is a chip that can be programmed to run custom code such as image segmentation and other image processing. An FPGA is capable of processing camera data with less latency than CPU or GPU approaches because it is parallelisable and has virtually no memory transfer cost. Using an FPGA-based solution would also mean that higher resolution cameras could be used without introducing additional overhead in the image processing stages.

8.2.2 Working area limitations

The blue-screen around the perimeter of the HOG table that facilitates background subtraction also limits the size of the table. Users need to be able to reach over it to perform gestures on the table surface. Techniques such as stereo imaging have the potential to facilitate 3D reconstruction without the blue-screen. Using such a technique the HOG table could support a significantly larger display surface and users would not be hindered by reaching over a high perimeter.

8.2.3 Situated HOG

The research presented in this dissertation has focused on large area communication. HOG objects are typically presented to the immersed user as larger than life objects. However, presenting the HOG table user's hand at a 1:1 scale has other potential applications. For example, a remote expert could provide instructional assistance for interacting with complex objects such as pipes and valves on an offshore oil rig. In this example, an immersed user on the oil rig would be able to see the hand of the remote expert as it demonstrates the manner in which the items can be interacted, without being restricted to a specific viewpoint. This is an interesting research problem as there are a number of questions to answer such as:

- How can the immersed user specify the object requiring assistance?
- How can HOG objects be registered to a particular object in an untracked environment?

This research idea has been labelled as Situated HOG since the life sized HOG objects need to be situated by the immersed user to show interaction approaches relative to the objects.

8.2.4 Mesh

The current approach used in this research does not explicitly generate a mesh of the reconstructed objects. Using a technique that does generate a mesh introduces a number of interesting possibilities. Collaborative modelling of outdoor structures would be possible. The HOG table user could provide a base object for carving against using the modelling tools built into Tinmith. Another possibility is manipulation of remote objects by the HOG table user. Using collision detection-based techniques the HOG table user would be able to user their hands to manipulate the position and orientation of existing HOG objects.

Generating a textured mesh could also potentially reduce the network traffic. Currently, processed images and computed geometry are transmitted over the network for every camera. Therefore, the quality of the reconstruction is limited to the available network bandwidth. But, if a textured mesh is generated on the HOG table side of the network, the number of cameras and the quality of the camera images would not have such as significant impact on network traffic to the rendering station.

8.2.5 Visualisation of collected information

The immersed user, particularly an outdoor AR Tinmith user, has the ability to collect large amounts of data such as pictures, videos and 3D textured models of structures outdoors. This information is potentially useful for people in a command and control center as it may help them gain a better understanding of the environment. Representing this information in an easy to access manner for the indoor users is a challenging problem. As seen in the user study in Chapter 6, imagery from the immersed user's viewpoint affected the HOG table users communication approach. However, many HOG users did not look at this imagery often as it took their focus away from the table. This is the type of issue that will need to be addressed to convey useful information to the HOG table user without distracting them from their current task.

8.3 Final remarks

God-like interaction is not currently a feasible solution for many applications. There are a number of technical issues that limit its immediate acceptance. While outdoor AR technology is evolving at a rapid pace, it is still not practical for many of the domains discussed in this dissertation. The HMDs are large and heavy, as are the computer systems required to run the graphics applications for them. Network bandwidth is very low, reliability and coverage are also problems that need to be addressed.

I believe that in the near future the technology to practically support the interactions techniques discussed in this dissertation will be commonly available. The technology to facilitate this interaction will likely fit into a pocket while also being robust and reliable. My hope is that the topics discussed in this dissertation help future researchers rethink traditional approaches to remote communication by leveraging the exciting possibilities that mixed reality and 3D reconstruction technologies provide.