

The God-like Interaction Framework: tools and techniques for communicating in mixed-space collaboration

Research Thesis for the Degree of Doctor of Philosophy

Aaron Stafford

Principle Supervisor:
Associate Supervisor:

Wayne Piekarski
Bruce H. Thomas



University of
South Australia

Wearable Computer Lab
School of Computer and Information Science
Division of Information Technology, Engineering,
and the Environment
University of South Australia

Declaration

I declare that:

- this thesis presents work carried out by myself and does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university;
- to the best of my knowledge it does not contain any materials previously published or written by another person except where due reference is made in the text; and all substantive contributions by others to the work presented, including jointly authored publications, is clearly acknowledged.

Aaron Stafford
Adelaide, November 2008

Wayne Piekarski
Adelaide, November 2008

Bruce H. Thomas
Adelaide, November 2008

Abstract

This dissertation presents techniques for remote communication of situational and navigational information. Outdoor augmented reality is a technology that facilitates viewing and generating geospatially registered virtual objects outdoors. In the future, emergency workers, law enforcers and military personnel may benefit from using outdoor augmented reality to help describe situational information to assist with safe and efficient completion of tasks. Situational information is constantly changing, what was true a minute ago may no longer be the case. These workers currently rely on a central command centre to keep them up to date with the larger picture.

This dissertation describes a new approach to communication of navigational and situational information between people in command centres, and people working in the field using outdoor augmented reality. This dissertation introduces the god-like interaction metaphor that describes a new approach to remote communication. The metaphor utilises a person's current perception about how an interventionist god influences the world through the hand of god. Command centre personnel visualise the remote workers on a large display, showing their position and orientation with respect to the real world. The command centre personnel use hand gestures and tangible prop interaction to provide navigation and situational information. A 3D capture system above a display captures the gestures and interactions, which are then presented to the remote field worker through the outdoor augmented reality system. To remote workers, the animated interactions appear to originate from the sky. Using this approach, command centre personnel can use their hands to point, make gestures, and represent objects and numbers. Tangible props can be used as virtual landmarks, virtual rally points and area of interest markers. Post-It notes can be used to create persistent information without the need for traditional input devices. To facilitate this form of interaction a 3D reconstruction tabletop display system was developed. The system has a number of cameras around the outside looking in that capture the interaction above a display surface.

3D capture systems have previously been developed to work over high-bandwidth reliable networks, but this system extends a well know video-based rendering algorithm to enable transmission of 3D content over an unreliable low-bandwidth network.

The practicalities and limitations of the god-like interaction techniques are explored in two studies. The first examines outdoor AR experiences, revealing problems with users estimating distance, particularly when the scale of the virtual objects is altered. However, participants are able to recognise objects and interpret intentions. The second study examines the comparative usefulness of gestures for providing visual cues for collaborative navigation. The results show that gesture-based cues are as effective as using a mouse to generate the cues. However, using gestures reduces the cognitive load on participants and facilitates experimentation of ways to convey intention, which is not possible using a mouse-based approach.

Acknowledgements

First, I need to thank my supervisors Wayne Piekarski and Bruce Thomas. There is a fable floating around the Internet about a rabbit who is writing a dissertation titled, “The Superiority of Rabbits over Foxes and Wolves”. It is an incredibly gutsy topic for a rabbit, particularly when confronted by so many foxes and wolves. However, the rabbit was smart enough to choose a lion for his supervisor and so was never in danger even when the foxes and wolves threatened to devour him. The moral of the fable is that the dissertation title doesn’t matter, the subject doesn’t matter, the research doesn’t matter, all that matters is who your supervisor is. I was lucky enough to have two lions.

I would never have had the opportunity to write this dissertation if it were not for Prabhu Manyem and Chris Stekete. Both were instrumental in facilitating me the wonderful opportunity of undertaking a PhD in the School of Computer and Information Science at the University of South Australia. I would like to thank them both for their belief in me.

Next up are the members of the WCL, an incredibly talented and highly motivated crew. Benjamin Close, Peter Hutterer, Ross Smith, Benjamin Avery, Aaron Toney and Mark Rebane shared an amazing journey with me. The WCL is an inspiring place to work, I have so many fond memories of my time there. I look forward to working on all sorts of crazy projects with this crew in the future.

There are a number of other academics I need to thank for their advice, guidance, time, patience, honesty and friendship over the years: Andrew McDowell, Georg Grossman, Vivian Nguyen, Puja Rawal, Sasa Baskarada, Grant Wigley, Stewart von Itzstein, Frank Fursenko, Phil Lock, Sue Tyerman, Ramesh Raskar, Roland Angst, David Kearney, Rudi Vernik, and Damien Bright.

As important as my supervisors were to guiding me down the road to submission, I would never have started or finished this dissertation without the love and support of my family and friends. Rebecca Witt—how lucky I was to fall in love with another academic. There are few people that understand the sacrifices and

commitment required to complete a PhD. Your love and support helped me to stay focused and sane. I would also like to thank my mum Mary, dad Peter, sister Naomi, brother-in-law Peter, the little'ns Jemma, Alicia, Vanessa and Samuel, and of course nana, for helping to shape who I am today.

Abbreviations and Definitions

AR Augmented Reality.

CCD Charge-Coupled Device.

FPGA Field-programmable gate array.

FOV Field of View.

GPS Global Positioning System.

GPU Graphics Processing Unit.

HOG Hand of god.

HOG table The 3D capture tabletop display.

HOG user A person operating the HOG table.

HOG object A 3D reconstructed virtual object generated using the capture technology of the HOG table.

HMD Head Mounted Display.

RGB Red, green and blue. A common colour space used in many graphics applications.

VBR Video-Based Rendering.

VR Virtual Reality.

Working volume The area on the HOG table in which all cameras can see.

Contents

1	Introduction	1
1.1	Problem statement	3
1.1.1	Research questions	4
1.1.2	Research goals	5
1.2	Thesis statement	5
1.3	Contributions	6
1.4	Dissertation structure	7
2	Background	9
2.1	Remote collaboration	9
2.1.1	Remote sensing	10
2.1.2	Remote instruction	12
2.2	Mixed reality	14
2.2.1	The real world	14
2.2.2	Virtual reality	14
2.2.3	Augmented reality	16
2.2.3.1	Outdoor augmented reality	18
2.2.3.2	Spatial augmented reality	21
2.2.4	Virtualized reality	22
2.2.5	Collaboration in mixed reality	22
2.2.6	Navigation	24
2.2.7	Interaction	26
2.2.8	Limitations	27
2.3	3D reconstruction	28
2.4	Tabletop displays	30
2.4.1	Interaction	31
2.5	Summary	32

3	God-like interaction	33
3.1	Representations of god in popular culture	34
3.2	The god-like interaction metaphor	36
3.3	God-like interaction techniques	37
3.3.1	Computer mediated techniques	38
3.3.1.1	Hand gestures	38
3.3.1.2	Props	41
3.3.1.3	Construction materials	45
3.3.1.4	Augmented Post-It notes	45
3.3.1.5	HOG on a WIM	46
3.3.2	Computer assisted techniques	49
3.3.2.1	Panning and zooming	49
3.3.2.2	Snapshot	50
3.3.2.3	Immersed reconstruction manipulation	51
3.4	Summary	54
4	Implementation	55
4.1	HOG table	55
4.1.1	Display	56
4.1.1.1	Panning and zooming	58
4.1.1.2	Projected display	58
4.1.1.3	LCD	59
4.1.2	3D capture	61
4.1.2.1	Working volume	62
4.1.2.2	Image processing	65
4.1.3	Limitations	67
4.2	3D reconstruction	69
4.2.1	High bandwidth reliable network	69
4.2.2	Low bandwidth unreliable network	71
4.2.2.1	Data compression	71
4.2.2.2	Handling lost packets	71
4.3	Visualisation systems	72
4.3.1	VR	73
4.3.2	Outdoor AR	73
4.4	Summary	74

5	Evaluation of god-like techniques	75
5.1	Evaluation	76
5.1.1	Task 1 - Recognition	77
5.1.1.1	Results	78
5.1.2	Task 2 - Interpretation of intent	80
5.1.2.1	Results	80
5.1.3	Task 3 - Judging distance	80
5.1.3.1	Results	82
5.1.4	Task 4 - Locating objects	84
5.1.4.1	Results	84
5.1.5	Task 5 - Questionnaire	86
5.1.5.1	Results	87
5.2	Summary	91
6	Collaborative navigation study	95
6.1	Hypotheses	96
6.2	Pilot study	96
6.2.1	Results	99
6.3	User study	100
6.3.1	Results	100
6.3.1.1	Task completion times	101
6.3.1.2	Effect of starting orientation	105
6.3.1.3	Effect of number of doors	107
6.3.1.4	Analysis of words used	110
6.3.1.5	Questionnaire results	112
6.3.1.6	Fatigue	113
6.4	Follow up study	115
6.5	Summary	117
7	Expert evaluation of the god-like interaction framework	119
7.1	Evaluation	119
7.1.1	General configuration	120
7.1.2	Virtual Landmarks	120
7.1.3	Pointing	121
7.1.4	Persistent Information	122
7.1.5	Photo-realism	122
7.1.6	Shoulder-to-shoulder collaboration	123

7.1.7	Symbology	124
7.1.8	Object recognition	124
7.1.9	Unconstrained interaction	125
7.2	Discussion	126
7.3	Conclusion	127
8	Conclusions	129
8.1	Contributions	129
8.1.1	Observations	130
8.1.2	Hardware	130
8.1.3	Software	130
8.2	Future work	130
8.2.1	Visual hull	131
8.2.2	Working area limitations	131
8.2.3	Situated HOG	132
8.2.4	Mesh	132
8.2.5	Visualisation of collected information	133
8.3	Final remarks	133
A	Outdoor user evaluation	135
A.1	Script	135
A.2	Questionnaire	147
B	User Study	151
B.1	Tabletop questionnaire	151
B.2	VR questionnaire	156
C	Low level implementation details	161
D	Camera quality	163
	References	166

List of Figures

1.1	A command center.	2
1.2	Tinmith - the outdoor AR system.	4
2.1	Spatially positioned remote collaborators.	10
2.2	3D teleconference.	11
2.3	The remote surrogate.	11
2.4	Netman.	13
2.5	Virtuality continuum.	14
2.6	An example of a VR system.	15
2.7	Outdoor AR.	17
2.8	Spatial augmented reality.	21
2.9	Shoulder-to-shoulder collaboration.	31
3.1	The hand of Ned	35
3.2	Playing a god in Black and White	36
3.3	Greek god metaphor in a virtual world	37
3.4	The HOG table	39
3.5	Outdoor AR view of a 3D reconstructed hand	39
3.6	Gestures and interactions are seen as animated sequences	40
3.7	VR view of a 3D reconstructed hand indicating left	40
3.8	3D reconstruction of a thumbs up gesture	41
3.9	Rock, paper, scissors.	42
3.10	Soft drink can navigation aid	42
3.11	Jar of Vegemite as virtual rally point	43
3.12	Flavoured milk representing an area to avoid	43
3.13	Prefabricated props example	44
3.14	Modelling structures with LEGO bricks	45
3.15	Augmented Post-It notes for persistent information	46
3.16	HOG on a WIM	47

3.17	HOG on a WIM techniques.	48
3.18	Third person view of snapshotted objects	51
3.19	Tinmith user relocated object.	51
3.20	Relocation of a cat	53
3.21	Wire-frame outline of 3D reconstruction.	53
4.1	HOG table designs and display images.	56
4.2	Detailed system diagram.	57
4.3	Second generation HOG table.	60
4.4	Levels of illumination possible with the display.	60
4.5	Camera mounts close up	61
4.6	Side view of capture volume of a single camera.	63
4.7	Working volume	63
4.8	Side view of working volume.	63
4.9	Possible configuration for increasing working volume.	64
4.10	Stages of image processing.	66
4.11	Good alignment of objects in reconstruction space.	68
4.12	Good alignment of objects in reconstruction space.	68
4.13	Data flow in a reliable network.	70
4.14	VR system and VR view.	72
4.15	Outdoor AR system and outdoor AR view.	73
5.1	Outdoor evaluation area.	76
5.2	Reconstructions shown to participants for identification.	78
5.3	Percentage of correctly identified objects.	79
5.4	Gestures for interpretation of intention	80
5.5	Six hands for participants to judge the distance to.	81
5.6	Summary of participants' ability to estimate distance.	82
5.7	Average times to locate hand.	85
5.8	Locating the hand against different backgrounds.	86
5.9	Awareness of weight of technology responses.	88
5.10	Field of view adequacy responses.	88
5.11	Backpack comfort responses.	89
5.12	Discomfort responses.	89
5.13	Backpack discomfort overlaid on backpack wearer.	91
5.14	Awareness of technology responses.	92
6.1	A view of the round virtual room.	97

6.2	The view of the virtual room on the HOG table.	98
6.3	A 3D reconstructed hand in the virtual world.	99
6.4	Travel time comparisons.	101
6.5	HOG table participants' ranking of efficiency.	103
6.6	Immersed participants' ranking of efficiency.	104
6.7	Starting orientation affect on audio condition completion times.	105
6.8	Starting orientation affect on mouse condition completion times.	106
6.9	Starting orientation affect on gesture condition completion times.	107
6.10	Number of doors affect on audio condition completion times.	108
6.11	Number of doors affect on mouse condition completion times.	108
6.12	Number of doors affect on gesture condition completion times.	109
6.13	Relationship between the number of doors and orientation to exit.	110
6.14	Average number of words per condition.	111
6.15	Immersed participants' questionnaire responses.	112
6.16	HOG table participants' questionnaire responses.	113
6.17	HOG table participant's responses to pain experienced.	114
6.18	HOG table participants' discomfort overlaid.	114
6.19	Comparison of times between audio and no audio	115
6.20	Comparison of the ease of attention grabbing	116
6.21	Concentration of HOG table participants.	118
C.1	HandOfGod object as a black box.	162
D.1	Reconstruction comparison with different quality cameras.	163