SPATIAL TASK PERFORMANCE IN VIRTUAL GEOGRAPHIC ENVIRONMENTS

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Adam John Rousell MSc (Leicester)
Department of Geography
University of Leicester

2012

Spatial Task Performance in Virtual Geographic Environments

Adam J. Rousell

It is well documented that within Virtual Environments performance in cognitive tasks is diminished, and with the continued use of such environments to train people in various skillsets it is important that this problem be addressed. In this thesis, two areas of spatial cognition are addressed: navigation and distance estimation. Unlike many previous studies, the experiments conducted here are in a large-scale virtual rural environment which poses problems due to the large distances involved and the unrestricted movement of people through it. A virtual representation of the Sorbas region in Spain was produced using Blueberry3D, VegaPrime and ArcMap.

Attempts to improve performance were made by the display of information to the user: an overview map to aid in distance estimations; and geo-located 'factoids', or info-marks, to aid in navigation. Analysis was also performed to extract rural environment features that could fall into the classifications of the Urban Image Theory, and a novel visio-analytic approach conducted to analyse track log data collected from the navigation task.

Results indicate that neither of the two tools implemented had much effect on user performance. However, a key finding was that the use of both quantitative and qualitative analysis is important in such research, as although quantitative analysis indicated only some significant results, the qualitative analysis highlighted that when the tools were presented users felt far more confident in their results. The visio-analytical approach adopted proved to be extremely useful in identifying performance characteristics that would have been missed by using quantitative analysis alone.

Acknowledgements

The experience of writing this thesis has been a long journey into the unknown, and as with any lengthy voyage there have been wrong turns and getting a bit lost. But the same as with the best trips, there have been new faces, new experiences and above all, that sense of satisfaction when you reach your destination and look back on the route you have taken. Throughout the whole experience, I have had great people travelling with me who have pointed me in the right direction when I've gone off course, and rode 'shotgun' with me to make the journey truly enjoyable.

Special thanks go to Dr Claire Jarvis, my supervisor, who has helped me throughout my PhD studies (and before) and helped me cope with the inevitable melt-downs (both mentally and technologically).

Also, appreciation goes to Prof. Chris Brunsdon who provided much needed assistance with the statistical side of the study.

Thank you also goes to the University of Leicester for providing funding for this study, and for being my home during the daylight hours (and sometimes even night time ones)!

Finally, a big thank you to all my friends, family and colleagues who have come on this journey with me, whether it be for a short ride or even from before it started, as without them I would not be here now.

"Do not go where the path may lead, go instead where there is no path and leave a trail." - *Ralph Waldo Emerson*

Table of Contents

1 Introduction 1

2 Literature Review 4

2.1 Virtual Reality and Transfer of Knowledge 4

2.2 Distance Perception 5

2.3 Perceiving and Navigating an Environment 16

2.4 Digitally Assisted Localisation 42

2.5 Visualisation of Movements 46

3 Rationale 54

4 Creation of the Virtual Environment 63

4.1 Site Selection 63

4.2 Data Collection and Preparation 65

4.3 Modelling 73

4.4 Blueberry3D Environment 75

4.5 Land Cover 79

4.6 Vega Prime 80

4.7 Research Laboratory Setup 82

4.8 Participant Selection 83

5 Part I: Observational Distance Estimation in VR 85

5.1 Introduction 85

5.2 Methodology 87

5.3 Results ... 106

5.4 Findings ... 121

5.5 Summary and Conclusions ... 124

5.6 Further Study ... 126

6 Part II: The Effects of Info-Marks on Performance in a Route Tracing Exercise ... 127

6.1 Introduction ... 127

6.2 Methodology ... 130

6.3 Results ... 168

6.4 Findings ... 201

6.5 Summary and Conclusions ... 207

6.6 Further Study ... 210

7 Part III: Visualisation of Track Data ... 213

7.1 Introduction ... 213

7.2 Data Extraction ... 214

7.3 Individual Analysis ... 220

7.4 Merging of Data ... 226

7.5 3D Display of Merged Dataset ... 231

7.6 Data Visualisation Against the VE ... 243

7.7 General Interpretation of Trial Results ... 252

7.8 Dataset Comparison ... 257

7.9 Summary and Advancements 267

8 Conclusions and Discussion 273

8.1 Summary of Findings 273

8.2 Discussion 279

8.3 Possible Modifications 282

8.4 Further Study 284

8.5 General Summary 287

9 Appendices 290

9.1 Appendix 1: Replicated Friedman Test Code 291

9.2 Appendix 2: Info-marks 294

9.3 Appendix 3: Track log Data 301

References 304

List of Tables

Table 1: Levels of data abstraction 3

Table 2: Depth cues 7

Table 3: Manhattan navigation feature types 41

Table 4: Spatial size of previous studies 57

Table 5: Types of environments used in navigation tests 61

Table 6: Data used in Blueberry3D model 66

Table 7: Distance estimation raw data 109

Table 8: Distance estimation trial over and under estimates (+ signifies over and - signifies under) 109

Table 9: Absolute relative errors for distance estimation task 111

Table 10: Absolute relative errors for transfer of knowledge task 113

Table 11: Estimation features 120

Table 12: Metrics used 156

Table 13: Track metric distributions 169

Table 14: KS test results 170

Table 15: Bartlett test results 171

Table 16: F-test results 171

Table 17: Area metric t-test results 172

Table 18: Track length metric t-test results 173

Table 19: Distance from end point metric t-test results 174

Table 20: Combined metric t-test results 175

Table 21: Travel duration metric t-test results 176

Table 22: Average speed metric t-test results 177

Table 23 : In-trial feedback (percentages) 182

Table 24: In-trial feedback classifications 183

Table 25: Ease and accuracy scoring 185

Table 26: F- test results (p-values) for ease and accuracy responses 186

Table 27: Ease of task Mann-Whitney test results 186

Table 28: Accuracy Mann-Whitney test results 187

Table 29: Use of info-marks 189

Table 30: "Rural Image Theory" responses 191

Table 31: Improvements to the environment 197

Table 32: Spearman rank test results of qualitative and quantitative metrics 198

Table 33: Dataset classification boundaries 227

Table 34: Data features 234

Table 35: Jenks classification boundaries 258

List of Figures

Figure 1: Optical convergence a) diagram of angle change; b) plot showing angle against distance 13

Figure 2: Distance cues 16

Figure 3: Navigation Routes in Manhattan 39

Figure 4: Space-time prism a) area of affect b) possible location between start and end points 53

Figure 5: DEM 67

Figure 6: Road and river network 69

Figure 7: Power lines and boundary features 70

Figure 8: Buildings 71

Figure 9: Land cover 73

Figure 10: 3D models created a) church, b) country house, c) swimming pool, d) cross 74

Figure 11: LOD modelling 77

Figure 12: Terrain classifications 79

Figure 13: Layout of research laboratory 82

Figure 14: Texture scaling 94

Figure 15: Map scaling 99

Figure 16: Trial locations 102

Figure 17: Boxplot of distance estimation absolute relative errors 112

Figure 18: Boxplot of transfer of knowledge absolute relative errors 113

Figure 19: Boxplot of distance estimation errors (by location) 114

Figure 20: Boxplot of transfer of knowledge errors (by location) 115

Figure 21: Info-mark locations 133

Figure 22: Info-mark dataset (left - raster, right - vector) 135

Figure 23: PDA info-mark delivery 136

Figure 24: ASCII file structure 139

Figure 25: Scene from the environment showing an info-mark display 140

Figure 26: Compass display 141

Figure 27: GPS sentence structure 144

Figure 28: Checksum code method 145

Figure 29: Training ground elevation (left) and land-use (right) 148

Figure 30: Training ground 148

Figure 31: Map given to users 151

Figure 32: Calculation of area metric 157

Figure 33: Extreme individual metric values 158

Figure 34: Area metric boxplot 172

Figure 35: Track length metric boxplot 173

Figure 36: Distance from end point boxplot 174

Figure 37: Combined metric boxplot 175

Figure 38: Travel duration metric boxplot 176

Figure 39: Average speed metric boxplot 177

Figure 40: Qualitative against quantitative data analysis 200

Figure 41: Map of UIT features 205

Figure 42: Track logs 215

Figure 43: Track counts 216

Figure 44: Clustering 218

Figure 45: Congestion 220

Figure 46: Track counts and environmental features (central region)........................ 222

Figure 47: Regions of high clustering.. 224

Figure 48: Central congestion region.. 226

Figure 49: Classified congestion and track count datasets 228

Figure 50: Merged dataset.. 230

Figure 51: IDW diagram of clustering 231

Figure 52: 3D visualisation of data 232

Figure 53: Data features: A) Mountain range, B) Volcano, C) Ridge, D) Trench, E) Plateaux, F) Valley.. 236

Figure 54: Blue valley and ridge area of interest.. 237

Figure 55: Blue valley comparison (coloured stars match locations in each environment) 238

Figure 56: High congestion and track count area of interest 240

Figure 57: Pre-turn ridge area of interest.. 241

Figure 58: Clustering between info-mark delivery methods A) no info-mark. B) onscreen info-marks, C) PDA info-marks.. 242

Figure 59: First turning point area of interest 243

Figure 60: Data visualisation and environment features 245

Figure 61: Blue valley environment comparison 247

Figure 62: Ridge and mountains environment comparison 248

Figure 63: Track convergence environment comparison 249

Figure 64: Mountain wall environment comparison 250

Figure 65: Dark mountains environment comparison 252

Figure 66: Central region info-mark delivery method comparison a) No info-marks; b) onscreen info-marks; c) PDA info-marks 260

Figure 67: End point region info-mark delivery method comparison; a) No info-marks; b) onscreen info-marks; c) PDA info-marks 262

Figure 68: Start region info-mark delivery method comparison; a) no info-marks; b) onscreen info-marks; c) PDA info-marks 264

Figure 69: Global view info-mark delivery method comparison; a) no info-marks; b) onscreen info-marks; c) PDA info-marks 266

Figure 70: Track count as elevation 270

Figure 71: Congestion as elevation 271

Map data (roads, rivers, elevation etc.) in figures 5, 6, 7, 8, 9, 15, 16, 21, 31, 41, 46, 47, 48, 54, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69 from the Spatial Data Infrastructure of Andalusia (Junta de Andalucía).

List of Abbreviations

ArcMap - GIS software developed by Esri

Blueberry3D - Fractal terrain generator software from Bionatics.

Bluetooth - Wireless data transmission protocol.

Creator - 3D modeling software developed by Presagis.

DEM - Digital Elevation Model – data identifying elevation of the ground surface.

GIS - Geographic Information Science/System – software and technologies used for manipulating and displaying geographic data.

GPS - Global Positioning System – System which uses satellites orbiting the Earth to provide a position on the Earths surface.

GUI - Graphical User Interface.

IL - Interaction Locus – area around an object inside a VE within which information is presented to the user (2007).

Info-Marks - Geolocated pieces of information presented to the user via an electronic device

LOD - Levels of Detail – representations of varying quality and detail for the same object.

LynX - GUI providing simple settings for VegaPrime systems.

MScape	- Software developed by Hewlett Packard to display information to users of a PDA dependent on a variety of sensory inputs.
NMEA	- National Marine Electronics Association – a protocol for transmitting GPS data between devices.
PDA	- Personal Digital Assistant – A hand held electronic device used for storing and editing files and information.
R	- Statistical package developed by the R Foundation for Statistical Computing.
UIT	- Urban Image Theory – theory created by Lynch (1960) whereby features in an urban environment are split into five categories based on people's perception and interaction with them.
VegaPrime	- C++ libraries for graphics and 3D environment display/interaction developed by Multigen-Paradigm.
VE	- Virtual Environment.
VR	- Virtual Reality.

1 Introduction

A growing number of entertainment systems, from cinema to hand-held game consoles, are now employing 3D projection systems, and with 3D films becoming increasingly popular the continuation of this shift seems set to develop further. In general, mass-market use of technologies results in manufacturing efficiencies and cheaper prices, setting into motion a cycle in which the use of 3D environments may well also increase. This poses problems since it is already known that performance in cognitive tasks tend to diminish in such environments (Ardito et al. 2007, Clément, Lathan & Lockerd 2008, Darken & Sibert 1996, Willemsen et al. 2008). Therefore ways of improving performance in elements such as navigation and distance perception are needed. After all, how much use is an expensive (in both time and money) 3D world if we cannot perform as well in it as the real world around us? If we are uncomfortable in a Virtual Environment (VE), we know that it becomes less immersive and so in turn may become less useful (Ardito et al. 2007, Taylor 2003). In terms of the static (vs. Interactive) 3D entertainment industry, cognitive navigation tasks may be less key but in a research and training setting where the 3D technology is used to display 3D training worlds, such tasks may well be integral to the experience and so methods of improving them for these instances would be highly beneficial.

This study investigates how the presentation of spatially rooted information can have an effect on cognitive tasks including navigation and distance estimation performed within a VE.

The literature review (Chapter 2) presented in this thesis investigates aspects including Virtual Reality (VR), distance estimation techniques, navigation methods, data visualisation, and the perception of environments. Particular focus is paid to the role of VR on training and assessment, cues used to estimate distances, particular cognitive methods required for navigation tasks (including cognitive maps), visualisation methods that can be used for displaying movement data, and the Urban Image Theory (UIT).

Based on findings from the literature review, the rationale (Chapter 3) sets out the research questions which this thesis tackles. By identifying the gaps in the literature regarding distance estimation, navigation performance (in particular landmarks) and visualisation techniques, questions emerge relating to how we might improve observational distance estimations in VR and whether the use of info-marks (geo-located pieces of information) might assist in a task (route-tracing). In examining how information can be gleaned from the data resulting from the many routes traced by research participants, a novel visual-analytic approach for exploring track data was developed. The detailed methods, results and discussions for these three themes are developed in Chapters 5, 6 and 7. Meanwhile in Chapter 4, the general methods used to develop the base environment in which research experiments are outlined: the area, data sets, software and modelled environment.

The common theme running throughout the thesis is the presentation of information and how this alters performance in cognitive tasks or aids in the analysis of data. In the first two investigations user trials are undertaken to monitor the effects of direct presentation of spatially rooted data on tasks frequently performed in both real and virtual environments: distance estimation and navigation. Navigation inherently requires the consideration of distance estimation, an important conceptual link. The final investigation takes the idea of presenting data to the user as a means of improving performance and directly uses visio-analytic methods to analyse data. Overall the thesis uses the visualisation of three forms of data (Table 1):

- Display of observable data.
- Presentation of hidden data.
- Visualising abstract data.

Display/Presentation method	**Information being portrayed**	**Space**
Overview Map	Visible features	Perceptual space
Info-marks	Hidden/nondescript feature	Contextual space
Track log visualisation	Data features	Conceptual space

Table 1: Levels of data abstraction

Moving down the table the abstraction of data increases, beginning with the presentation of data relating to features that are readily visible in the scene. This is followed by data referring to features that are present in the environment but could be missed and so are seen as 'hidden', and then finishing with data that are not present in the environment but are rather collected through users' interaction with it. Thus, the thesis works across several data spaces: perceptual, contextual and conceptual.

The thesis concludes by bringing work from all three spaces together in one summary and reflects on aspects for further study that arise from the research, alongside a critical evaluation of the research as a whole.

2 Literature Review

2.1 Virtual Reality and Transfer of Knowledge

VR has many applications such as interactive data analysis (Kellogg et al. 2008), psychological experiments (Witmer et al. 1996), and training for real world activities (Feudner et al. 2009, Lucas et al. 2008, Schoor et al. 2006, Watanuki & Kojima 2006). Systems have been developed to replicate real world areas (i.e. Nurminen (2008) and Witmer et al. (1996)), environments based on real world situations (i.e. Schoor et al. (2006)), and abstract environments which would not be visually possible in the real world (i.e. Biggs et al. (2008)). Each application has a different level of immersion (the sense of being 'within' the environment) depending on the hardware system used. One useful aspect in the use of VEs is that knowledge gained within them can transfer to real world applications. Training tasks previously implemented include those in the fields of medical treatment (Feudner et al. 2009, Schoor et al. 2006), industrial training (Lucas et al. 2008, Watanuki & Kojima 2006), and medical rehabilitation (Brooks & Rose 2003, Lam et al. 2006). In all of these applications, the user has entered a VE and then performed tasks within it such as navigating a large building as in the case of the study by Grant & Magee (1998). Indeed, training tasks previously implemented also include those in the field of navigation (Farrell et al. 2003, Grant & Magee 1998, Jansen-Osmann & Berendt 2002, Lathrop & Kaiser 2005). Lathrop & Kaiser (2005) used a virtual representation of the International Space Station as a means of investigating whether the level of immersion (feeling of being within the environment) affects navigational performance. Jansen-Osmann & Berendt (2002) made use of a desktop VE to assess how the number of turns in a route can affect the perceived length of the route. Farrell et al. (2003) investigated how well a route learned within a VE can transfer to a real world location, and Grant & Magee (1998) investigated how the use of pro-prioceptive feedback (use of a treadmill rather than a joystick to convey movement) altered performance when transferring spatial knowledge from a VE to its real world counterpart.

Of particular relevance to this thesis, it has been demonstrated that navigation processes performed within VEs use similar methods to those performed in real world environments (Aoki et al. 2008, Biggs, Fischer & Nitsche 2008, Bosco et al. 2008, Sturz et al. 2009) and that navigational skills learnt can transfer from the virtual to the real (Giudice et al. 2010).

2.2 Distance Perception

It is well documented that distance is frequently underestimated in both Real Environments (REs) (Proffitt 2006) and VEs (Bodenheimer et al. 2007, Clément, Lathan & Lockerd 2008, Interrante et al. 2008, Willemsen & Gooch 2002, Willemsen et al. 2008), and that 'real world' judgments are more accurate (Willemsen & Gooch 2002). Alexander et al. (2003) note that these differences between real and virtual depth perceptions arise as a result of a combination of the characteristics of the tracking and rendering system, the display being used, and the visual deficits of the user.

There are many cues that are used to perceive distance and depths of objects. These include motion parallax[1] (Clément et al. 2008, Uehira et al. 2007, Willemsen et al.

[1] Motion parallax – The apparent movement of features at different speeds depending on their distance from the viewer whilst the viewer is moving in parallel to them

2008, Witmer & Kline 1998), stereopsis[2] (Clément et al. 2008, Hendrix & Barfield 1995, Hubona et al. 1999, Witmer & Kline 1998), shadowing (Hasenfratz et al. 2003, Hendrix & Barfield 1995, Hubona et al. 1999, Jurgens et al. 2006), textures (Clément et al. 2008, Hendrix & Barfield 1995, Hubona et al. 1999, Willemsen et al. 2008, Witmer & Kline 1998), and differing sizes of objects within the environment (Hendrix & Barfield 1995, Hubona et al. 1999, Interrante et al. 2008, Willemsen et al. 2008, Witmer & Kline 1998). It is worth noting however that some cues have dominance over others (Hubona et al. 1999), and that cues such as motion parallax and stereopsis in particular are combined with others to aid in distance and size perception (Rauschecker et al. 2006). These terms, and their meanings, are introduced in more detail in the sub-sections below. In addition, some cues are limited in terms of the maximum distance they can be used for (i.e. accommodation and convergence can only be used up to 3 metres whereas size can be useful up until the horizon (Cutting & Vishton 1995)). Particular depth regions have been described by Cutting & Vishton (1995) as personal space (within arm's reach), action space (2-30m) and vista space (30m +). The cues themselves can also be split into binocular or monocular cues depending on whether both eyes are required or not for the cue to work effectively (see Table 2).

An observation worth taking into account however is that most research in respect of distance perception takes a psychological or physiological approach; it focuses on what allows us to perceive distance as opposed to a more practical approach that reflects upon what might be done to improve distance perception. For real world applications this is a shortcoming, since it is important that a person can perceive distances

[2] Stereopsis – The effect of 3D vision arising from each eye viewing the same scene from a slightly different location

accurately. Lappin et al. (2006) for example report the case where underestimating distance from the observer (pilot) to the runway can result in the pilot believing that they are travelling slower than they actually are, potentially contributing to collisions with terminals. In a geographic context, being able to determine distance between the observer and features within a large scale environment where physical measures cannot be easily taken would be very useful for navigation (Witmer & Kline 1998). It would also be valuable for more analytical tasks such as creating sketch maps, quickly providing information about the environment, or undertaking geophysical analysis.

Method	Monocular/binocular cue
Shadowing and Drop Lines	Monocular
Motion Parallax	Monocular
Textures	Monocular
Size	Monocular
Stereopsis	Binocular
Accommodation	Binocular
Convergence	Binocular

Table 2: Depth cues

2.2.1 Shadowing and Drop Lines

Shadowing has a major effect in VEs, not only adding realism but also in providing information about the spatial location of an object (Hasenfratz et al. 2003). However, most experiments where shadowing is investigated consist of trials where the objects are not in direct contact with the ground surface (i.e. Hendrix & Barfield (1995), Jurgens et al. (2006) and Kersten et al. (1997)). Is it the case that cast shadows are only of use for perceiving distance when the object and shadow receiver are not in contact?

There appears some discrepancy between experiments as to the effectiveness of shadowing in comparison to other cues (e.g. Hendrix & Barfield (1995) and Jurgens et al. (2006)). Hendrix & Barfield (1995) performed a study to compare effects of shadows and drop lines (a vertical line drawn from the object to the ground surface) in respect to assessing distance and elevation of objects within a computer generated abstract scene. They found that although the use of shadows did improve the error values in respect to positioning tasks, their effect was minimal in comparison to the effects of drop lines. Jurgens et al. (2006) on the other hand noted that a variety of projection methods (including shadows and drop lines) proved to be as useful as each other. They determined from qualitative feedback that a 'cast circle' representation (one where a circle is drawn on the ground below the object which decreased in size as the object grew closer to the surface) was better than the others (standard shadow, drop line, numerical representation), although from the statistical analysis of results no one method was better than the others. The study by Jurgens et al. (2006) revolved around actions performed within personal space and action space (0-30m) in that the experiments involved the placing of a stake held by the test subject onto a feature on the ground.

Hubona et al. (1999) also performed experiments using shadowing and found that although they did aid in positioning tasks, this was only the case if a stereoscopic display was not used and the shadows came from a single light source. If multiple light sources were used, a single object created multiple shadows which negated any benefits of adding shadows to the scene and results were similar to those where no shadowing was present. This is an important aspect to note as computational costs of including shadows within VEs can be expensive (Hasenfratz et al. 2003) and so if they are only of benefit when stereoscopic vision is not implemented, then adding them to a stereoscopic environment would not be of any benefit in terms of user performance.

Another feature that occurs in the natural world that is closely related to shadowing is reflection (Jones & Bertamini 2007). Jones & Bertamini (2007) performed experiments to determine the effects of presenting reflections created by planar mirrors within a virtual scene, and found that these reflections increase the accuracy of relative size and distance estimations between two objects when they both have reflections present. However, they also note that although this improvement is apparent, reflections scarcely appear in the natural environment in comparison to shadows, and thus may not be a particularly useful cue.

2.2.2 Motion Parallax

Motion parallax is the effect whereby when objects move laterally across a person's field of view, closer objects appear to move faster than more distant objects (Witmer & Kline 1998). It is an absolute depth cue meaning that it can be used to determine explicit distances (in known values such as metres and kilometres) between the observer and object (Willemsen et al. 2008).

However, unlike many other depth cues, motion parallax requires that the perceiver is moving (Clément et al. 2008, Uehira et al. 2007). As such, this method would not be any use in improving distance perception in tasks where a user is stationary. Although, as motion parallax is a monocular depth cue (Clément et al. 2008), it can be used for larger distances (including over 30m as noted by Uehira et al. (2007)).

2.2.3 Textures

Textures on an object whose distance from the observer is to be assessed, or on other features within an environment (such as walls), can be used to aid in distance perception (Clément et al. 2008, Hubona et al. 1999, Willemsen et al. 2008). This arises since textures appear smaller on more distant objects. In addition a repeating texture can also be used as a means of counting distance. Textures can be viewed as a relative cue in that they can be used to say that an object is twice as far away as another object, but you cannot explicitly imply distance in metres or kilometres (Willemsen et al. 2008).

It seems that in some cases textures are only effective when used with other depth cues. For example, in their trials Hendrix & Barfield (1995) found that texturing was only effective if there was a drop-line connecting the object to the textured surface. However, Witmer & Kline (1998) notes that if a plain flat wall was approached so that it filled the user's view, without a texture it would be very difficult to judge how far away you are from the wall itself. Texture, in the latter case, is therefore the key. There is a predominant difference between the placement of textures in these two studies however. In the study by Hendrix & Barfield (1995) the texture was not on the object itself but a ground surface, whereas the texture discussed by Witmer & Kline (1998) is on the object whose distance is to be estimated. Therefore it is clear that the use of a texture as a distance estimating cue is dependent on the object that it is actually on.

Something worth taking into account when using ground surface textures to aid in distance perception is that if the texture is interrupted (either by an object or change in texture) then the distance to an object past this has been found to be underestimated (Wu et al. 2007a). Therefore ground textures may not be the best method for use in real world environments (or virtual representations of these) since it is likely that there will be a number of texture changes between the observer and the object whose distance away is being judged. To judge a distance accurately, the linear perspective of the surface is also important, as shown by Wu et al. (2007b). Here linear perspective refers to the convergence of parallel lines on the horizon. In this study, textures were created where lines were not parallel but in fact converged. If these lines are made to converge prematurely and the horizon is not visible, then a perceived horizon created is lower than that of the actual horizon resulting in an underestimate of distance (Wu et al. 2007b). This arises because a horizon can be estimated using linear perspective, a discontinuity of optic flow (the breaking of continuous objects due to occlusion by other features), and the use of compressions in ground textures (Bertamini et al. 1998). This then results in a smaller angle between the perceived horizon and target, making it appear farther away. If, on the other hand, the horizon is visible, then a premature convergence gives the impression of an incline on the surface (Wu et al. 2007b).

2.2.4 Size

The apparent size of an object can affect its perceived distance, or vice versa. Emmert’s law suggests that an object’s perceived size is proportional to its perceived distance. The Moon for example, appears larger when it is close to the horizon; this is because it is perceived as being further away on account of the distance cues provided by the ground surface (Kaufman & Kaufman 2000). Also, the size of familiar objects can be used to aid in judging of sizes of other objects such as architecture plans and models using people and trees (Interrante et al. 2008). There is however evidence to suggest that providing such familiar objects within a VE does not aid in a more accurate perception of distances (Interrante et al. 2008) when the objects are used as guides. On the other hand, Nakamizo & Imamura (2004) performed comparisons of the effectiveness of Emmert's law in both virtual and real environments, finding that it holds in both. However, they note that in the VE it only holds up to 20m as opposed to 24+m in real environments. The difference between these two studies was that Nakamizo & Imamura (2004), unlike Interrante et al. (2008), judged perceived size as opposed to the distance. Additionally, Interrante et al. (2008) used the objects as a guide to determining the size of other features (a room) whereas Nakamizo & Imamura's (2004) investigation was conducted in relation to the objects themselves.

Another study by Predebon & Woolley (1994) investigates the effects of 'familiar size' on depth perception. This is where the relative apparent size of two objects can be used to determine their depths in respect to each other (i.e. a closer object appears larger than a more distant one), although this relationship may be better described as relative size. The main finding from the investigation was that this familiar size relationship was superseded by the information gained from the motion parallax cue, suggesting that it is not an effective cue for determining depth.

2.2.5 Stereopsis, Accommodation and Convergence

Stereopsis is the effect of each eye seeing a slightly different image causing a stereo 'depth' effect (Witmer & Kline 1998). As it requires the use of both eyes to work, it is one of the few binocular depth cues (Clément et al. 2008), alongside cues such as accommodation and convergence. Accomodation refers to the focusing of the eyes to objects at differing distances, and convergence is the rotation of the eyes to look at a specific object (Liu et al. 2008). There is a general consensus in the literature that stereopsis is possibly the most powerful and dominant of depth cues (Hubona et al. 1999). Grant & Magee (1998) suggest that stereopsis may support the distance estimation component of path integration – a method of navigation and orientation (using distance travelled and acceleration/velocity to determine location which is discussed in a later section).

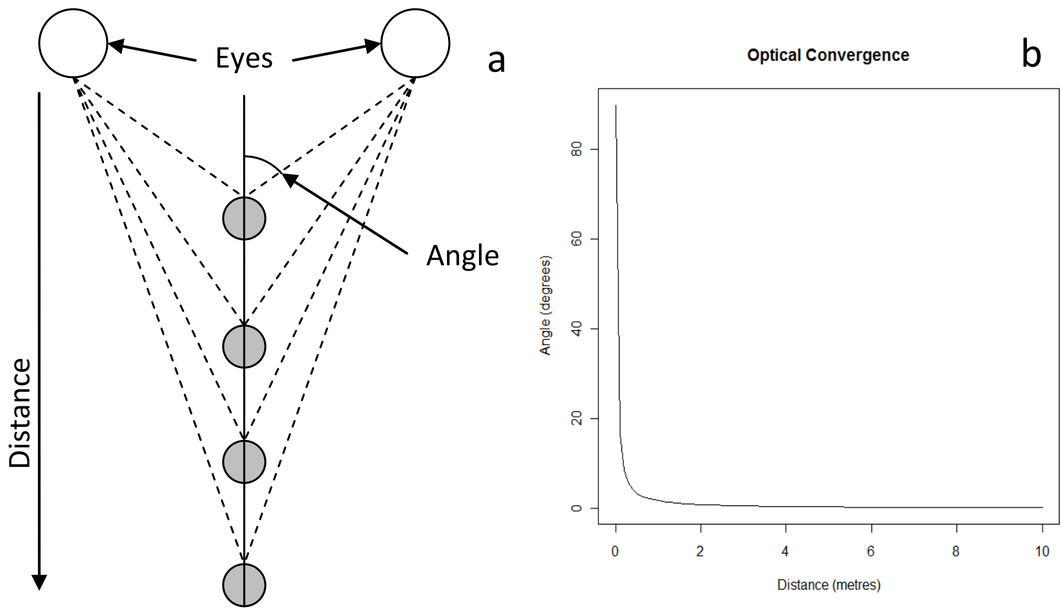


Figure 1: Optical convergence a) diagram of angle change; b) plot showing angle against distance

It is worth noting that as an object becomes more distant, stereopsis is a less effective means of judging distance (Hendrix & Barfield 1995), although Allison et al. (2009) believe that stereopsis should be effective up to a range of greater than a kilometre. Additionally, as a distance increases we see that the difference in angle between eyes to look at the same feature (convergence) decreases dramatically (see Figure 1a). The convergence method is thus unlikely to be significant at geographical scales (see Figure 1b).

2.2.6 Traversed Distance

An additional aspect of distance perception is that of ‘traversed’ distances (or distances that have been travelled), as opposed to ‘observational’ distances (or perceived/viewed distances). In a study by Witmer & Kline (1998), different methods of movement within a VE were tested to see which resulted in the least amount of estimated distance error. They found, perhaps surprisingly, that the increased proprioceptive feedback (feedback relating to motion or position of parts of the body) offered by a treadmill as a movement method did not improve the estimation of traversed distance as much as first thought. However, they also discovered that a key factor in determining traversed distances is in fact the speed at which you move. When used in conjunction with navigation and orientation tasks, it has also been found that the use of a treadmill did not aid in orientation tasks (Grant & Magee 1998). However, Grant & Magee’s (1998) study does improve the transfer of knowledge from VEs to real world counterparts, as can be seen by a decreased amount of distance travelled to reach a goal.

Crompton & Brown (2006) also investigated the relationship between environmental structure and perceived traversed distance. They found that in environments with more turns, intersections and features of interest, people overestimate the distance travelled by an average of 3 times. This compares with an average of 1.6 times for a fairly straight and regular road within a city environment. The other interesting finding of this investigation is that the traversed distances were over-estimated, as opposed to the usual under-estimation found when distances between objects are judged.

Exploring more recent technology, Suma et al. (2007) performed a study into using tracking systems as opposed to a treadmill for navigating within a VE, and found that participants using a real walking technique completed tasks faster and with fewer collisions than those without. This links with an earlier study by Zanbaka et al. (2005) which showed that using real-walking methods increases cognition and 'presence'[3] within the environment. Although these studies do not analyse any aspect of traversed distance, they do highlight that using non abstract navigation methods does increase presence within an environment.

Finally, estimating a distance between two places that are not necessarily in view (cognitive distance) uses different processes to estimating distances between objects in view (psychophysical distance) (Crompton & Brown 2006). This is an important factor when investigating navigation and orientation, since different approaches for aiding in distance perception must be used between orientation tasks and navigation tasks in cognitive versus psychophysical situations.

2.2.7 Combined Cues

In general, it is not just one cue that we use to determine distances but a combination between multiple cues. For example, in Figure 2 we can see a simple landscape containing two trees and a path. In this example, the ground is a continual texture-less surface (i.e. snow cover) and so the ground texture aspect cannot be used. However, the edges of the path converge as it approaches the horizon, and the shadows formed

[3] Presence can be seen as the sense of 'being within' the environment rather than simply observing it

from the trees spatially connect them to it. Measuring the height of a tree could be problematic as, when looking up, perspective can distort our perceptions; however, we could easily determine the width of the pathway. Therefore, if we know that the path is a constant 4 metres in width we can make assumptions as to how far away the path is crossed by the shadow of the more distant tree based on the linear perspective effect.

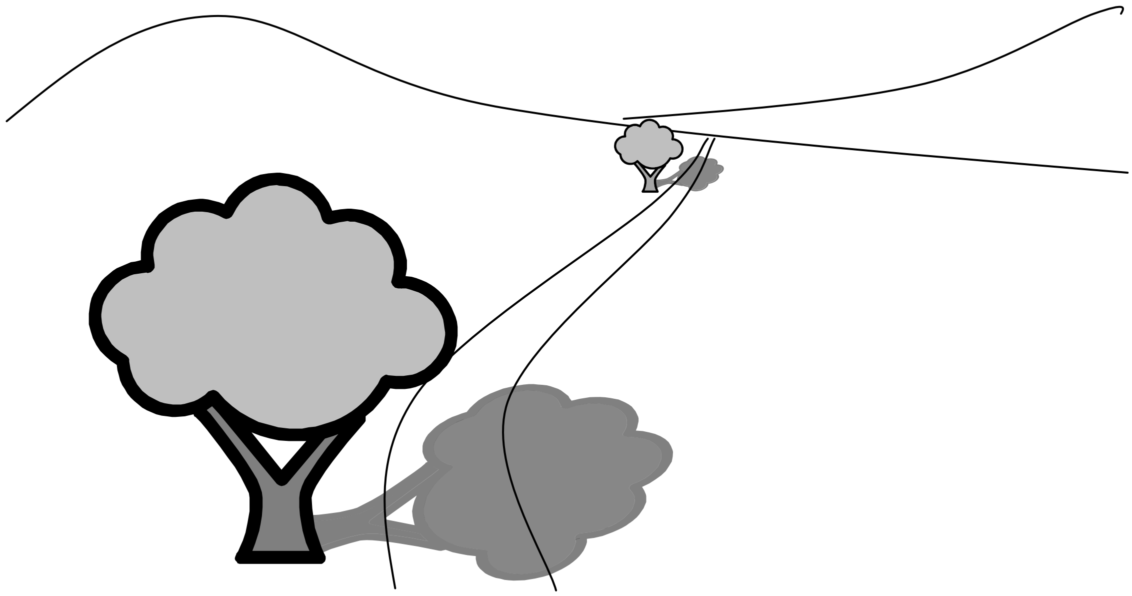


Figure 2: Distance cues

2.3 Perceiving and Navigating an Environment

When we look at the environment around us we see different features and how we use (and in turn possibly perceive) these features varies depending on the task that is to be performed within that environment. In this section, how we mentally store our perception of the environment around us in the form of cognitive maps is reviewed. In turn, how these maps in our heads are used for navigational tasks is investigated. Additionally, this section elaborates upon how the social structures of the environment can effect what is perceived. In connection with this last theme, a major study into how urban environments are structured and understood (the UIT) is also discussed with an emphasis on how it can be used to understand navigation.

2.3.1 Cognitive Maps

Siegel & White (1975) described that to form a spatial representation[4] three forms of knowledge are constructed which are: information about landmarks, information about the routes between landmarks, and information about the general configuration of the environment. It is now widely accepted that when learning any environment (either virtual or real), three forms of knowledge are created: landmark, route and survey (Abu-Obeid 1998, Janzen et al. 2001, Ruddle & Péruch 2004, Sjölinder et al. 2005). It is also generally understood that these three forms of knowledge are constructed at the same time (Hurlebaus et al. 2008, Ruddle & Péruch 2004); earlier work suggested a sequential acquisition of knowledge from landmark leading to route leading to survey knowledge (Lathrop & Kaiser 2005, Parush & Berman 2004, Siegel & White 1975). In fact, the construction of survey knowledge can be obtained without a full construction of route knowledge (Janzen et al. 2001, Sadeghian et al. 2006). To construct this spatial representation, it is believed that locomotion within the environment is an almost essential condition (Siegel & White 1975). It should be noted that there is some debate as to whether the landmark-route-survey model can account for all aspects of navigation, as in some instance the development of survey knowledge develops almost instantaneously (Darken et al. 1998, Zhang 2008).

If an environment contains more important (or easily recognisable) elements for building a basic cognitive framework map, then the environment is likely to be more

[4] The term “cognitive maps” was not chosen for their study owing to it implying a map like structure, which is not necessarily the case

‘legible’ to the observer (Huynh et al. 2008). Here, both landmarks and structure are considered key components when building the framework for a particular area (Elvins et al. 1997).

This section continues by considering landmark, route and survey knowledge in detail.

2.3.1.1 Landmark Knowledge

Landmark knowledge is the information obtained about landmarks within an environment and is *“acquired through the acquisition of information about distinctive environmental features.”* (Abu-Obeid 1998 pp 160). Landmarks themselves are *“unique configurations of perceptual events (patterns). They identify a specific geographical location.”* (Siegel & White 1975 pp 23). Janzen et al. (2001) describe landmark knowledge as “*a few unconnected landmarks [that] are stored in memory.”* [pp 150]. It is important to note that landmarks themselves do not need to be singular distinct objects but can be (for example) road junctions (Ruddle & Péruch 2004, Siegel & White 1975), and can also be seen as meaningful events in time (Siegel & White 1975). The key aspect of landmarks is that they are geographically rooted in some way, whether this be through the configuration of road intersections, the uniqueness of a structure, or the spatial aspect of a memorable event.

2.3.1.2 Route Knowledge

Route knowledge can be seen as being *“developed through linearly connecting discrete environmental features”* (Abu-Obeid 1998 pp 160) and refers to the ability to navigate between locations (Ruddle & Péruch 2004). It is egocentric (Morganti et al. 2007), that is, it is centred around the person itself as opposed to other aspects of the environment – it emerges from the eyes of the user with themselves at the centre. It is the integration of landmarks into paths and sequences used for navigation (Wu et al. 2009).

2.3.1.3 Survey (Configurational) Knowledge

Survey (or configuration) knowledge is a generally considered to be the 'overview' of the environment that allows for flexible navigation between two points (Hurlebaus et al. 2008), although the definition does vary somewhat between studies. Siegel & White (1975) define configuration knowledge as being *"constituted by gestalt knowledge"* [pp 24]; in other words, the environment is seen as a unified whole. This amounts to considerably more than a simple compilation of route and landmark knowledge. Other studies such as Janzen et al. (2001) indicate that survey knowledge is the integration of multiple routes, while Morganti et al. (2007) define it as the state achieved when *"distant places are linked together to form an integrated global overview of the entire environment."* [pp 1984]. Although the definitions of survey knowledge are not always the same, one aspect that is constant is that it is a larger scale representation of spatial layouts and relationships within the environment. After the acquisition of survey knowledge, several cognitive determinations can be undertaken (Witmer et al. 2002):

- Distances between landmarks and destinations can be calculated.
- Directions to destinations can be calculated.
- Shortcuts within routes can be engineered.

For all three types of knowledge, various factors influence the construction of spatial knowledge regarding an area. For example, Meilinger & Knauff (2008) found that when constructing spatial knowledge, either the use of verbal directions or the use of a map resulted in the same overall effectiveness for building a cognitive map. This outcome is believed to arise since the routes acquired from the maps were being stored mentally as a series of directions; these series of directions are similar to those given verbally. It should be noted however that in the field of learning, Hawk & Shah (2007) discuss different learning style models which describe how different people learn. In these models different methods are employed which cater to specific aspects of learning (i.e. visual, oral, reflective) and it is shown that a model on its own (of those that they investigated) cannot provide for all learning styles. Therefore, this same concept of different methods of learning being employed by different people could account for some of the differences in how survey knowledge (and other spatial knowledge) is acquired between individual people. Witmer et al. (2002) note that the means by which knowledge is obtained *can* affect how useful it is. Survey knowledge for example is gained faster by the use of an overview map, but the quality of this information is then superseded by survey knowledge gained from extended navigation within the environment. Additionally, the gaining of survey knowledge can be impeded by using proprioceptive feedback (Ruddle & Péruch 2004) which is the sensing of movement and position of parts of the body.

In sum, the effectiveness of approach does depend on the type of knowledge being obtained. Additionally, the characteristics of the learner will also come into play.

For any one individual learner, educational models suggest that the preferred learning style (e.g. visual, auditory, kinaesthetic) may strongly influence the effectiveness of different agencies for building cognitive maps (Hawk & Shah 2007). It could be hypothesised for example that Meilinger & Knauff's (2008) results are a good reflection of an overall position across different 'types' of learners, with individual variation. Visual learners may have an initial advantage when developing survey knowledge as they could find it easier to ‘see’ the environment as a whole.

It has also been identified that older people tend to have deficits in the construction and/or use of cognitive maps (Iaria et al. 2009, Sturz et al. 2009). However, there is some uncertainty as to whether there is a gradual diminishing or a more discrete cut-off point in this process. In a study by Head & Isom (2010) with regards to effects of age on cognitive maps, for example, age bands of 18-22 were used for younger and 56-83 for older meaning that a large proportion of the population (23 to 55) was not assessed. It could in fact be that there is a gradual diminishing of performance from an age that is present in this un-assessed age group.

A further aspect relating to the learning of an environment is the spatial literacy associated with the person learning it. Jarvis (2011 pp 294) describe spatial literacy as “*a state reached through the practice of spatial thinking*” and suggest that the development of it is an iterative process involving the linking of spatial abilities, strategies and knowledge into spatial thinking, which with continued use forms spatial literacy. Woollett & Maguire (2010) performed a study using London taxi drivers who can be seen as expert navigators of the London area. When asked to produce a sketch map of a new urban environment that they had just been subjected to via videos of routes through it, the taxi drivers outperformed people who were novices in navigation. However, when the same experiment was performed using a modified London environment (mixture of features from London and Bath, UK) it was found that the taxi drivers found the integration of the new features particularly difficult. Therefore it is clear that being a good navigator in one environment has a positive effect when learning a new one, but a person who has a large amount of spatial knowledge about a specific location may have difficulties if the layout of that location is altered. As Jarvis (2011) notes, both spatial strategies and knowledge combine to build spatial literacy over time.

2.3.2 Navigation and Orientation

The literature regarding navigation and landmarks is extensive and includes aspects such as neurophysiology (Koene et al. 2003), robotic simulation (Smith et al. 2006), surgical navigation (Stüdeli 2009) and even positioning in outer space (Yang et al. 2010). The literature reviewed in this study focuses on the cognitive navigation and orientation aspects of the genre, and the way in which different types of landmarks aid in this process.

Navigation itself is essential in any environment that requires movement over large distances (Darken & Peterson 2001). Navigation can be described as:

> *"getting from one point to another within a given environment or geographical area"* (Parush & Berman 2004 pp 376)

Navigation itself can be split into two aspects, wayfinding and travel (Dodiya & Alexandrov 2008, Sadeghian et al. 2006). Wayfinding is defined as the determination and following of a route, and travel understood as the locomotion along that route to reach the destination (Dodiya & Alexandrov 2008).

It is important to note that when orientation is discussed in regard to navigation we are not simply looking at where something is positioned, but rather a concept, as highlighted in the quote below from Parush & Berman (2004).

> *"[Orientation is] the ability to know one's location within the environment and the relative location of other elements, and to continually update this knowledge"* (Parush & Berman 2004 pp 376)

The processes both of navigation and orientation can be closely related to the forming and use of cognitive maps. Developing a cognitive map is an important core task that is performed within both real world and virtual environments (Bosco et al. 2008, Burigat & Chittaro 2007, Chittaro et al. 2006, Ruddle & Péruch 2004). To navigate successfully we need to know our orientation (where we are) within the environment (Grant & Magee 1998, Harrower & Sheesley 2007, Kelly & Bischof 2008, Parush & Berman 2004). Finding the way between two locations is a task performed frequently by both humans and other animals (Hurlebaus et al. 2008). Orientation can be performed using different methods such as piloting (using landmarks and spatial knowledge) and path integration (monitoring velocities and direction whilst travelling) (Grant & Magee 1998). In fact Palermo et al. (2008) state that:

> *"... the use of environmental cognitive maps is essential for orientating since it allows individuals to reach any target location starting from any place and by following any route available within the environment"* (pp 249)

Just as in the case for the building of cognitive maps, individual differences in approach and ability have been highlighted within the literature. In Sturz et al.'s (2009) experiments for example, cognitive maps were not necessarily used during navigation. Strategies learned during training, as opposed to any constructed cognitive map, are also sometimes used. Studies have also found that increased age causes detrimental effects on both navigation and orientation (Iaria et al. 2009). It is also believed that males and females navigate differently with an overall suggestion that males tend to prefer distance travelled and direction based navigation, whereas women tend to prefer landmark based navigation (Mueller et al. 2008); both methods are equally valid forms of navigating within an environment. When a lack of spatial knowledge is available to a person, information can be passed on by another person to allow them to navigate successfully within it (Bidwell & Lueg 2004), for example giving verbal directions.

Methods to improve the quality and construction of users' cognitive maps should therefore have a positive effect on peoples' ability to navigate and orientate themselves within an environment. Zhang (2008) note that navigation can also be made easier if the environment is well structured using specific design principles including the UIT and pattern languages (discussed in later sections). On the other hand however, emphasising specific features in the environment to aid in navigational tasks may have a detrimental effect on creating a sense of presence within the environment (Steiner & Voruganti 2004). A key aspect of navigation is with respect to features that are easily identifiable in the environment that can be used to help identify location and destinations – features which are known as landmarks.

2.3.3 Landmarks

There is extensive research into landmarks and how they are defined and used. For an overview of what constitutes a feature being noted as a landmark, see Caduff & Timpf (2008) and Sorrows & Hirtle (1999).

Landmarks are a key aspect in navigation. This is the case whether they are used for verbal direction giving (i.e. Peters et al. (2010)) or for general orientation and identification of target locations (i.e. Pierce & Pausch (2004)). Landmarks located at decision making points along a route (i.e. turning points) are both seen as more important, and are recognized more quickly (Miller & Carlson 2011). They are equally essential for navigating in both real and virtual environments (Lazem & Sheta 2005). With specific reference to navigation in VEs, Vinson (1999) suggests that landmarks are distinctive visual features that can be used as reference points, both for navigation and orientation. In his study he sets out several guidelines with regards to the creation and implementation of landmarks within VEs. These guidelines are identified through literature review (although the study did not test the effectiveness of implementing the guidelines) and are split into four aspects: learning about an environment; landmark types and functions; landmark composition; and minimizing distortions in cognitive maps. The guidelines set out are as follows:

- Learning about the environment
    1. It is essential that the VE contains several landmarks
- Landmark type and functions
    2. Include all five types of landmarks (UIT) in your VE[5]
- Landmark composition
    3. Make landmarks distinctive
    4. Use concrete objects, not abstract ones, for landmarks
    5. Landmarks should be visible at all navigable scales

[5] By landmark in this instance, it is actually meant the five feature types from the Urban Image Theory: paths, edges, districts, nodes and landmarks.

6. A landmark must be easy to distinguish from nearby objects and other landmarks
7. The sides of the landmark must differ from each other
8. Landmark distinctiveness can be increased by placing other objects nearby
9. Landmarks must carry a common element to distinguish them, as a group, from data objects
10. Place landmarks on major paths and at path junctions

- Minimizing distortions in cognitive maps

11. Arrange paths and edges to form a grid
12. Align the landmarks' main axis with the path/edge grid's main axis
13. Align each landmark's main axis with those of other landmarks

As well as the aspects set out by Vinson (1999), the guidelines can also be split into two types, the first being the *appearance* of the landmarks (1-9) and the second being the *positioning* of landmarks (10-13). Obviously if the VE is meant to be a realistic representation of a location in the real world, then it is unlikely that changing the actual structure of the environment would be desirable. In that case it would be more beneficial to highlight specific features and objects already present so that they can become landmarks.

One of the main items addressed in this list of guidelines is the uniqueness of the landmarks. The guidelines from Vinson (1999) suggest that features that are landmarks should be different in some way from data objects in the environment; they should be easily distinguishable from other objects and landmarks. He also highlights that the landmark should look different depending on the direction in which it is approached (i.e. sides must differ). By making the landmarks unique it should become easier to orientate oneself. It is also suggested that distinctiveness can be enhanced by placing other objects nearby, as recorded by Alexander et al. (1977). Their book 'A Pattern Language' (Alexander et al. 1977), demonstrates how different types of patterns that form in environments alter the way in which people interact with them.

It is not just within physical spaces where landmarks are important and where typologies of landmarks have emerged. They are also significant in electronic or cyber-spaces (Sorrows & Hirtle 1999). In Sorrows & Hirtle's study, navigation in electronic space is with regards to movement around hypertext (web pages). Landmarks in this case can be items such as navigation bars, FAQ pages or the home index page. They also define three types of landmarks: visual, cognitive and structural. In the case of cognitive landmarks, they are defined as having a meaning which stands out, unlike the cognitive maps which are mental representations of space. Each of these types has a specific aspect that makes them landmark worthy, such as visual characteristics, meaning, and location, although the categories are not discrete and stronger landmarks will fall into all three categories (Sorrows & Hirtle 1999). Also, Sorrows & Hirtle (1999) note that the predominant type of landmark used will vary dependent on the type of navigational task. Travel to a new location is more likely to benefit from the use of cognitive and structural landmarks, whereas movement to a known place may use more visual and cognitive based landmarks.

Similar to the landmark types noted by Sorrows & Hirtle (1999), Caduff & Timpf (2008) describe the salience of landmarks (the prominence and distinctiveness of features) outside of cyber-space in terms of perceptual salience, cognitive salience, and contextual salience. Perceptual salience is how the landmark is perceived in the environment in terms of sensory input, cognitive salience is how the feature is seen in terms of the viewer's own experience and knowledge, and contextual salience is how much attention can be allocated to the identification and use of available landmarks which is altered by the navigational task and method of locomotion.

Turning back to geographical environments, not all landmarks will be visible at any one time. Thus the landmarks defining the goal location may not be visible from the navigator's current location (Smith et al. 2006). Pierce & Pausch (2004) assessed the benefits of providing visible landmarks at all times to users of a large scale VE; this required that the landmark being used was dependent on the scale of the area that it was representing (i.e. a tall building for a district and smaller buildings with more characteristics for a street). As opposed to showing all landmarks at all times, a hierarchical method was implemented. At the lowest level scaled replicas of each landmark were used; at greater distances, and so a higher level in the hierarchy, symbolic representations were used to show the area rather than the landmarks within it. Results indicated that providing such representations made traversing large virtual worlds more efficient. In addition, providing global landmarks aids in the determination of location within an environment (Rousell et al. 2008). In the study by Rousell et al. (2008) a large-scale VE was used with global landmarks being cone shaped structures in the sky above the location of villages and towns. Users were asked to plot on a map (showing the locations of the towns and features such as roads) their perceived location. When the markers were displayed users were considerably more accurate at determining their location. It was also found that providing information about distance from the user to the marker in the form of a graduating colour scale did not improve estimations against when plain coloured markers were used, although this could be due to the colour scale not being a suitable method of portraying distance.

Attempts have been made to identify landmarks automatically through computerised methods, with the aim of improving route guidance systems. Claramunt & Winter (2007) successfully used graph theory and space syntax as a means of identifying salient features based on Lynch's (1960) UIT (see next section) which could later be used as landmarks to aid in the generation of automated route directions. Elias (2003), adopting an approach echoing earlier pattern-language research, used a digital cadastral map to identify buildings that were unique amongst their surroundings yet might not necessarily be used as a landmark if were found somewhere else. Using such a method as seen in pattern language research, it is possible to not just identify the classic landmark structures (i.e. Empire State Building, Buckingham Palace) but also the ones which are more difficult to categorise (i.e. multiple buildings that form a unique layout). By extracting information from the map relating to aspects such as size, orientation, building density and distance from road and storing this information in a database, the buildings which could form landmarks for the particular navigation task could be extracted. Raubal & Winter (2002) also presented a study whereby landmarks for use in a navigation service were derived from a database using a series of properties based on the types of landmarks identified by Sorrows & Hirtle (1999). Importantly, Sorrows & Hirtle (1999) identify types which can be used to categorise landmarks in both geographical and hyper-spaces rather than purely geographical aspects. In the case of Raubal & Winter (2002), the types were used to identify geographical landmarks, although in a hyper-space context the same method could be employed to produce a navigation service that would direct a user through web pages. Each structure in the database used was assessed based on aspects such as its shape, historical importance and location and then statistical methods used to derive whether it should be used in the route giving instructions. A problem with this study however is in one of the definitions of the model used:

> *"Along edges no orientation action needs to take place. The traveller shall move from the start node to the end node (i.e., from a decision point to a decision point)."* (Raubal & Winter 2002 pp. 246)

Although this is likely the case in structured environments, other large scale environments such as rural areas may not have linear 'corridors' that ensure that between decision points travellers stay on the correct path. Therefore in such open areas, landmarks may need to be constantly highlighted along the edge to ensure that the traveller does not deviate from it.

Another aspect mentioned in the study by Raubal & Winter (2002) is the potential of *active* landmarks. These landmarks (such as electronic beacons) are used by mobile devices to provide navigational landmarks to users via software but are not explicitly experienced in the environment by the person. Therefore they cannot be used as reference points for human wayfinding (Raubal & Winter 2002). However, an aspect that was not addressed in that study was the potential to link active landmarks directly to features in the physical environment; for example, they might be used to highlight a structure or object that may have otherwise been overlooked. After being introduced to the user via the active landmark, the feature may then become a landmark itself.

The active landmark does not itself become the useful feature in a wayfinding task, but has the potential for human wayfinding.

From the varied topologies for landmarks and guidelines provided for the generation of landmarks in real and virtual spaces, the literature on this subject is broad. It is clear that what defines a landmark is not just a single aspect but a large number of properties that are not just physical and visual. Most recent research focuses on attempts to extract landmarks automatically from datasets, but as the following review sections suggest, there remains further work to be done examining how social components might be considered as part of the landmark research.

2.3.4 Social Structures

How features are perceived can be affected by aspects other than their physical structure and geographic location. When we look at environments, in particular urban areas, human interaction can have a dramatic effect on how we perceive and use specific locations and features which in turn may affect a specific features role as a landmark. Micarelli & Pizzioli (2008) state:

> *"landscapes can be conceived as complex phenomena, where human, natural life, and built environments are evolving as one, showing their different characteristics, contrasts and dynamics."* [pp 2]

This implies that the urban structure and the people who live there are inherently linked and 'evolving as one'. To perceive the urban environment around us suitably, we must also perceive the people who live there and the social interactions that occur. In the study by Shamsuddin & Ujang (2008) for example, the physical and social character of traditional shopping streets in Kuala Lumpur are investigated in relation to place attachment. The study confirmed what we might intuit; the identity of the streets is not just created by the physical structures found there but also the historical significance, density of people and their nature as socio-cultural strongholds. Even the movement and colour along them make one street distinctive when compared with another.

Whether or not a person likes a location or chimes with a particular sensory landmark can be influenced by many factors. If a person prefers social aspects such as culture and social exchanges, they may well identify more with urban environments than those who prefer greenery and tranquillity (Félonneau 2004). This is an important aspect to note, since Félonneau (2004) recognises that a person who is able to identify with and is attached to a city is more likely to filter out aspects such as dirtiness and disrespectful behavior than those who are not. This in turn could have a dramatic effect on what aspects of an environment a person has a sensory identification with. A person who is attached to an urban environment may not notice the sound of traffic, whereas someone who is attached to rural areas might.

It is not just personal experiences and preferences however that can affect how someone feels about a particular location. Savitch (2005) identifies that urban terrorism can have a drastic effect on how people perceive a particular location. Locations that have been attacked (or potential attack sites) have mental boundaries created around them which may turn a social magnet into a repelling force. In a study by Richardson & Jensen (2008), it emerged that the use of the platforms on the Bangkok Skytrain network in music videos invoked a sense of ‘progress and modernization’ and lent these locations a strong iconic factor. Favro (2006) attempted to identify a single object that can be used similarly to the Skytrain of Bangkok to iconicise ancient Rome. She found that in such a historically rich place, one single physical or abstract object cannot suitably identify it. From the attempted iconisation of Rome, and the symbolic importance of Bangkok's Skytrain, we see that the physical features alone may not be as important as what they represent. In the study by Pierce & Pausch (2004), it was iconic objects that were implemented to aid in navigation within larger scale virtual worlds. In that study, each area was depicted by a single visual landmark based on either a physical object within it or an artistic, symbolic, representation.

2.3.4.1 Urban Image Theory

A key study in the field of how we perceive and structure the environment around us is known as the Urban Image Theory, proposed by Lynch (1960). This theory describes how urban environments are structured with regards to human perception, and thus how the theory itself can be used to aid in the development of new areas, making them easier to navigate and better structurally defined. Although designed for real world places, other studies have suggested that it also has implications for virtual worlds (Modjeska & Chignell 2003); the theory has even been used as an aid to surgical wayfinding inside a human body (Stüdeli 2009). According to UIT, all features in an urban environment can be categorised into one of five groups:

- Paths – linear features that can be traversed (roads, pathways, railroads etc.).
- Edges – linear elements that cannot be traversed and can act as linear breaks between two areas (fences, walls, shorelines etc.).

- Districts – a two dimensional area that the observer mentally enters and has some defining characteristic (business district, Chinatown etc.).
- Nodes – points or locations within the environment which are of particular importance which can are entered by the observer (transportation hubs, road junctions, social communing areas etc.).
- Landmarks – point features that are unique and can be used to identify location (a statue, aesthetically pleasing building, a mountain etc.).

As can be seen, several of these features (e.g. paths and edges) are closely related to each other. However, the main conceptual difference between paths and edges is how people interact with them. On paper, they are both linear features, but they are very different in how they are used and perceived in the field. For example, the fence of a park is clearly an edge; it is linear, cannot be traversed and forms a distinct barrier. In this same way, a footpath through the park is clearly a path; it is a linear feature that can be traversed as a means of getting from one place to another. However, if we take the example of a main road it can be seen as both an edge and a path depending on who perceives it. For a motorist, the road is clearly a linear feature that can be travelled along. Pedestrians on the other hand may very well walk beside the road on the pavement but they do not (usually) walk on roads and they can be difficult to cross. Therefore the pedestrian may view the road as an edge.

Similarly, nodes and landmarks can both be shown on paper as point features. The main difference between them however is how they are used in the process of navigation. A cathedral within a city can be used as an example for this; if a person's navigational aim is to get to the cathedral as an end point or a stop off along a longer journey then this would be a node. This is because the cathedral is an important point in the navigational process. However, if the person uses the cathedral as a locational identifier then it can be seen as a landmark. It is still important in the environment as a whole, but the person doesn't actually enter it (in this navigational exercise anyway). Even if the cathedral is used as an explicit navigational aid such as in the directions "turn right at the cathedral", the cathedral itself is still a landmark as the physical location that the person turns (i.e. road junction) is the node.

The areal district features can be one of the more difficult feature types to define physically, as they tend to be more of a mentally defined region as opposed to a geographically defined one. In many cases, the mental region does indeed correlate with some physical area within the environment, but the boundary of the physical area is not what defines the district. In addition, a district must be easily recognisable from the inside so there must be some global characteristic about it that allows this recognition. These characteristics are endless; they could be some sort of architectural standard, the structural layout, the use of the area, or even defined by the people within it. More often than not, it is multiple features taken together that make the district recognisable (i.e. in a business district of a large city you would expect tall buildings and people wearing suits; in a dock area you would expect to hear the sound of boats and water, as well as different smells depending what goods are normally unloaded there). In addition to this, the boundaries of the districts can become somewhat 'fuzzy' depending on what defines the district. For example, if the district is surrounded by some form of edge feature, then the district boundary would generally be this discrete edge. However, if there is no distinctive edge (i.e. commercial and business districts) then the districts may bleed into one another and the boundary becomes much fuzzier.

In a general Urban Image that is used to describe the environment as a whole, it is likely that people will classify most features in a similar way. If these classifications are used to distinguish features in a navigational context then it is likely that the classification for a particular urban theory object may change between navigational tasks. To show how this occurs, two scenarios are now presented (see Figure 3 for a graphical representation of the routes). In the first scenario, there is a man who has just arrived at Grand Central Station in Manhattan, New York. He needs to make his way to the corner of Christopher Street and 7th Avenue in Greenwich Village to meet a friend, but he does not want to take the subway. In this instance, he may have directions such as:

- Leave Grand Central Station onto 42nd Street and turn right.
- Continue until you reach 5th Avenue junction where there is the library. Turn left onto 5th Avenue.
- Continue down 5th Avenue, passing the Empire State Building and crossing Broadway and entering Greenwich Village until you reach 8th Street. If you reach Washington Square you've gone too far.
- Turn right onto 8th Street and continue to the junction with Avenue of the Americas.
- At this junction, turn right and then turn left at the next junction which is Christopher Street.

- Carry on down Christopher Street until you come to the first junction, which is with 7th Avenue.

In this scenario, he starts at Grand Central Station and so this can be seen as a node. Next he is told to walk down 42nd Street which means that 42nd Street can be seen as a path (as can 5th Avenue, 8th Street, Avenue of the Americas and Christopher Street). He is told that he needs to turn left at the junction with 5th Avenue where there is a library. In this statement, the junction becomes a node as it is a turning point and the library can be seen as a landmark. Along 5th avenue he is also told that he will pass the Empire State Building and cross Broadway, and so both of these can be seen as landmarks as they are used to show that he is going in the right direction. Washington Square can also be seen as a landmark as it is also a feature that is used to identify whether they are going in the right direction, although this time it is used to show an incorrect decision. Also along this section of the route they are told they will enter Greenwich Village. As Greenwich Village is an area within Manhattan that has specific characteristics that set it apart from other areas, here it can be seen as a district. After making a couple of additional turns and junctions (nodes) they arrive at their destination which is the corner of Christopher Street and 7th Avenue.

In a second scenario, there is a female tourist who is staying at a hotel on 30th Street between Madison Avenue and Park Avenue. For her morning sightseeing she wants to visit the Empire State Building to go to the viewing level at the top, before going to Grand Central Station for some cheesecake. Her directions might be as follows:

- Leave the hotel and turn right onto 30th Street.

- Walk down 30th Street crossing Madison Avenue until you reach 5th Avenue. Turn right onto 5th Avenue.
- After 3 blocks you will find the Empire State Building on your left.
- Upon leaving the Empire State Building, turn left onto 5th Avenue and continue until you reach the junction with 42nd Street just after the Library. Turn right onto 42nd Street.
- After crossing Madison Avenue you will come to the junction with Park Avenue – Grand Central Station is on your left.

In the case of this person, the start point is the hotel and so this is the starting node. She is instructed to walk along 30th Street which is a path, crossing Madison Avenue (a landmark here) before reaching the junction with 5th Avenue (a node). The next major feature is the Empire State Building. In this scenario however, this is a target destination and so is a node rather than a landmark. After leaving the Empire State Building they are told to continue along 5th Avenue, turning right at the junction with 42nd Street (node) after the library (in this scenario also a landmark). After crossing Madison Avenue (landmark) they will find Grand Central Station on their left, which is the end destination so a node.

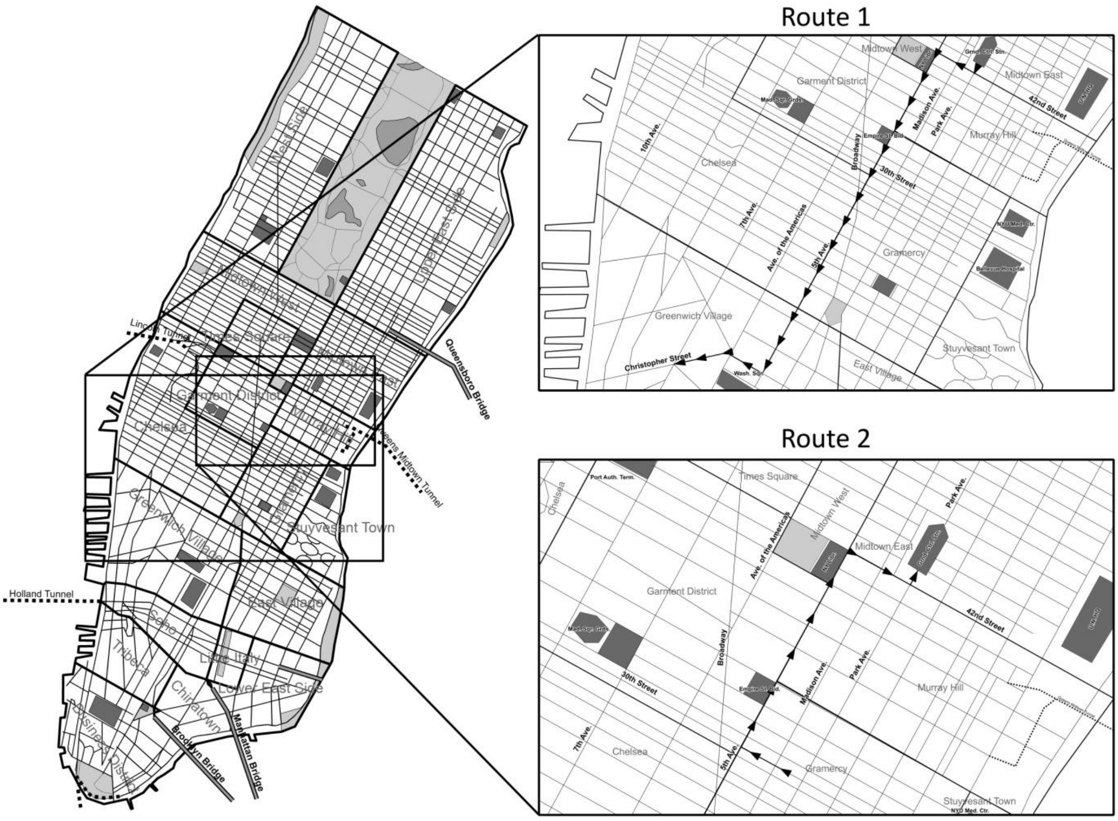


Figure 3: Navigation Routes in Manhattan

Here it is evident that features in one context have a different type than the same feature in another context. The primary example here is the Empire State Building which in the first scenario is a landmark as it is used to identify that the person is going in the correct direction. In the second scenario it is a node as it is a key destination along the route that the person enters. A list of different feature types used in each scenario can be seen in Table 3.

Type	Scenario 1	Scenario 2
Path	42nd Street 5th Avenue 8th Street Avenue of the Americas Christopher Street	30th Street 5th Avenue 42nd Street
Edge	-	-
District	Greenwich Village	-
Node	Grand Central Station Junction of 42nd & 5th Junction of 8th and 5th Junction of 8th & Avenue of the Americas Junction of Christopher St. and Avenue of the Americas Junction of Christopher St. and 7th	Hotel Junction of 30th & 5th Empire State Building Junction of 42nd and 5th Grand Central Station

Landmark	Library Empire State Building Broadway Washington Square	Madison Avenue Library

Table 3: Manhattan navigation feature types

The basic elements themselves can be used to describe specific features within an environment, but to describe the environment as a whole suitably it is the *relationships* between these elements that creates the form of the environment. These relationships can have an effect on the elements themselves. For instance, a small district may seem important when first considered, but if there is a large landmark within it then it is most likely that the landmark will take precedence and the district will be more or less forgotten about.

When we look at an Urban Image, it is normally a general image that we look at meaning that it is not defined by a specific task. Therefore, when an image is created, it is not specifically designed for the purpose of helping a tourist getting from one place to another; it is created by people who know the urban environment well. This means that specific features can be assigned a general type i.e. the Empire State Building would be a landmark and Madison Avenue a pathway if everyone in Manhattan was asked. In a study by Omer et al. (2005), residents of Tel Aviv were asked to draw on a map the dominant features they perceived in the Tel Aviv urban environment. If an element was drawn by more than one person, it was assigned a type and drawn on an aggregate Urban Image map. To further highlight the imageability (how many people would see them as part of an overall Urban Image) of features, they were re-drawn to highlight the frequency of appearance (0-25%, 25-50%, 50-75% and 75-100%) using visualisation techniques such as line weight and shading. By using this method, it is easier to generate a more generalised image for the urban environment that would be representative of the people living in that environment.

Another aspect of the UIT that has received little (if any) attention is how the classifications described can be fulfilled with features present in rural environments. In such areas, could the theory if taken more generally still be applied? As rural environments have a far lower concentration of highly distinctive urban features, what rural aspects might be used to replace them?

2.4 Digitally Assisted Localisation

In recent years mobile devices (i.e. PDAs, cameras, smart phones) have been gaining locative technologies including accelerometers, thermometers and GPS at increasing accuracies and decreasing costs (Counts & Smith 2007). In addition to this, new devices have ever increasing computing power and high-speed internet connectivity (Koh et al. 2009) and when coupled with web-browsing software being present in many mobile devices (Pires et al. 2010) an inevitable surge in systems integrating web-based technologies and geographical rooted information will be seen. In fact, websites such as Flickr provide facilities for everyday casual users to geo-tag their photographs after they are uploaded (Yaegashi & Yanai 2009), allowing the images to have a geographical grounding.

Location and Context-Based Services (LCBS) are services which have an inherent relationship between application and location or activity-contexts (Kühn 2004). Some applications of LCBS as described by Kühn (2004) are:

- *Intelligent navigation support* – use of constantly updating databases relating to traffic and car parking spaces to provide dynamic routing information for cars.
- *Transport logistics* – relationships between rail freight or the distribution of workpieces, tools and machines in a smart factory.
- *Rescue/emergency support* – coordination of actions between participating parties in the event of a catastrophic situation.
- *Automatic collision avoidance system* – use of 2D, 2.5D and 3D location awareness to avoid collision between fast moving vehicles and suddenly appearing objects.
- *Mobile working environment* – access points where data can be cached locally for use later on.
- *Spatial information systems* – communication between users as to determine spatial relationships (i.e. who is close, messaging to users within a specific area, identification of areas).
- *Location/context-based events* – triggering of events dependant on external conditions (i.e. altering the temperature of a room when someone enters/leaves, reservation of services for a delivery vehicle once it is within a certain distance).

The above examples all make use of specific geographic information as to where an actor is in relation to another actor or system. Although each uses this information for a different purpose, it is clear that the eventual outcome is dependent on some geographic factor. This geographical factor can be determined through several instruments and technologies (Kühn 2004) including:

- GPS,
- Cell-Of-Origin within a mobile phone system,
- Handset-/network-assisted triangulation within cellular phone systems,
- Active infra-red badges for indoor positioning,
- Ultra-sound based location determination,
- Electronic maps,
- Vehicle sensor data (acceleration, speed, bearing etc.),
- Direct user input.

Counts & Smith (2007) describe the use of mobile devices to generate “hyperties” which are a link between hyperlinks generated in computational media and the interactions people undertake with surrounding objects and their environment. The main example given is the use of locational data (geographic coordinates and accelerometer readings) alongside biological data (heart rate) to generate route logs generated from exercise activities (i.e. running, hiking, cycling). By collecting all this information and storing it in a web-based database system, people can view their own routes and those generated by others and make decisions about alternative routes or social activities. For example, a person could increase the difficulty of their route by merging and splicing sections of their own and other people’s routes to include more elevation changes or specific regions where a higher heart rate is generally present. Also a person's route could be compared to those of other people to see if there is anyone who does the same route and so may wish to jog together. This example highlights that the tying and presentation of data from multiple sources can be highly beneficial not only on a personal level but also a social one.

Other applications of LCBS include tourist information systems (Jeong et al. 2006, Pires et al. 2010) and student field studies (Field & O'Brien 2010). Field & O'Brien (2010) make use of Twitter posts (tweets) to generate mash-up maps of locations and key topics being discussed. With regards to a field trip, students were asked to post tweets using a specific hash-tag (a tag that denotes a particular topic, in the case of that study #malta09) along with GPS information to observe and discuss information about their surround area. These tweets were then extracted using the Twitter search API and plotted on a Google Map so that students could collaborate and discuss what was being found.

With the use of software such as Hewlett Packard's MSCape (see Stenton et al. (2007)), datasets of geo-tagged data can be quickly created and then used on GPS-enabled mobile devices to display information relevant to the location the user is currently in. This system is conceptually similar to the Interaction Locus (IL) described by Ardito et al. (2007) in which areas within a VE (as opposed to 'reality') are associated with external data such as text, visual, auditory and tactile information. The same concept of IL has been previously implemented by Fogli et al. (2003) where a PDA system was developed for use within a museum. In the case study provided, three ILs were used depending on user preferences, and relative information was presented to the users.

Information about prior interaction with the system was recorded, resulting in the user being made aware if they had viewed all features or revisited a previous one.

All of these studies show that the addition of a geographic aspect to datasets can have a highly beneficial effect on how the data are viewed and analysed. Counts & Smith (2007) and Field & O'Brien (2010) show how location based information (information that has a location attribute) collected from multiple users can be shared as a means of improving understanding of an area and altering interactions with that area. Ardito et al. (2007) and Fogli et al. (2003) on the other hand use the same concept of adding a geographic dimension to data but use it to determine when that piece of information should be displayed to a dynamically moving 'recipient'. Therefore it is clear that adding the geographic dimensions to data has a great benefit in multiple applications.

2.5 Visualisation of Movements

A common method of assessing how people navigate through data stored in information space is to record a user's actions and then use this information to identify particular patterns (Ieronutti et al. 2005). However, although this method is commonly used in the context of web sites and hypermedia, there are little in the way of analysis methods for analysing navigational performance and characteristics within VEs (Drachen & Canossa 2009a, Ieronutti et al. 2005). There are some methods being proposed (for example Börner & Penumarthy (2003), Chittaro et al. (2006), Chittaro & Venkataraman (2006), Hoobler et al. (2004) and Ieronutti et al. (2005)), although most of these visualisations rely on presenting the information to the user via a 2D representation, often only showing one particular aspect of the track log data.

In a system called VU-Flow (Visualisation of Users' Flow) (see Chittaro et al. (2006), Chittaro & Venkataraman (2006) and Ieronutti et al. (2005)), track logs are collected from multiple users and then analysed off-line (post-visit (Chittaro et al. 2006)) to identify several characteristics. Such analysis included the identification of:

- Areas where more time was spent.
- Areas that were travelled more or less.
- Areas that were seen more.
- Locations that were congested.
- The direction of flow of users.

One primary advantage of VU-Flow is that it is cross-application compatible. This means that it does not rely on data collected from a specific application. However, only one piece of information is presented to the user at any one time so an analyser can view only the flow direction or the congestion on a single map at any one time. To view both they would need to either have two displays or the same display but each map shown at a different time. For simple and small maps this is not a particular issue, but as the size and number of the maps increase, this method could become increasingly difficult to identify key areas. Hoobler et al. (2004) also visualise user movements (this time in a computer game) but use a vector based approach rather than the raster representations used in the VU-Flow system. In this case, multiple pieces of information can be shown at any one time, with a line being used to identify the path taken by a user which alters in thickness depending on how recent it was taken. Additionally, glyphs were placed on the visualisation to identify the current location of users as well as their job role (soldier/medic/engineer), team colour, and their current health. The paths of munitions fire were also displayed, as well as user field of view. Finally, a grid representation with varying colour and intensity was used to identify area control based on the team that was last in the grid cell. Hoobler et al.'s (2004) study shows that multiple pieces of information can be presented on a single display via the use of different icons and representations. A major issue with this form of representation however, is the number of users that can be adequately visualised. With increasing numbers of users, the required number of different visualisation techniques so as not to make the visualisation unreadable also increases (Börner & Penumarthy 2003).

In the study by Börner & Penumarthy (2003) a visualisation toolset was described to aid in the analysis of aspects such as user movements, chat analysis, and the spatio-temporal diffusion of groups and individuals. In the study, users of 3D browser systems where web pages are linked to 3-dimensional objects and are presented over the internet had movements and communications recorded for analysis. Alongside the tracks generated by the users, the toolset also visualised the information regarding the location of items such as teleports and where chats occurred, as well as showing the location of the 3D models with a colour scale used to represent their age.

Drachen & Canossa (2009a) used data relating to users position and actions within a computer game which was logged every one second. After collection, GIS software (ESRI's ArcGIS) was used as a means of analysing the number of deaths on a particular map area using the heatmap method, which is a raster based visualisation where the colour and/or intensity of cells changes depending on the value being measured. In the case of Drachen & Canossa's (2009a) study, the heatmap was used to highlight areas that had a high lethality in the gaming environment (locations where users were frequently 'killed'). The benefit of analysing such data is that areas that could possibly be too difficult (a high death rate) could be altered to make the area easier to survive.

An important aspect regarding the visualisation of data is with regards to the actual information that is collected from the environment. Drachen & Canossa (2009b) discuss the collection of data relating to gameplay metrics which are specific values relating to gaming. For example, in their study information about the users navigation (x,y,z coordinates taken every second), health (x,y,z coordinates where the player gets injured), crouch and cover (x,y,z coordinates where the player performed a crouch or snap-to-cover action which is a hallmark of the game being assessed), and the speed (information as to whether the user was standing still, walking, running etc.) were analysed. The primary goal of the study was to highlight that this information can be used to assess user navigation within the game and to determine whether players deviated from the intended path, as well as to identify areas that may be particularly difficult. They document several benefits of using gameplay metrics including:

- Providing highly detailed quantitative data.
- Supplementation of existing methods.
- Assisting in the location of problems.
- Providing progressive detail depending on the level of analysis.

Zanbaka et al. (2005) also used visualisation techniques to analyse user movement within a VE, via 'spaghetti plots' where all tracks were placed on top of each other alongside 'dwell data' (the percentage of time spent in each location). Again, the visualisations here are only shown one at a time, and the environment they are collected from is particularly small and restrictive in terms of movement allowed (the environment was a single room measuring 4.5 x 4.6 x 2.6 metres). Additionally, they note that the intersection of paths in the spaghetti plot highlight common intersections in the environment. It is unclear however as to what is actually being seen by these as they do not take into account the time on each path that intersections occur. For example, two paths crossing at a point where one person was in that location twenty seconds into their trial and the other ten minutes would not be particularly useful.

In some instances it may be of benefit to know where a person has not been as opposed to the paths that have been taken. Imbe et al. (2010) discuss a system known as MyGlobe where users can share their preferences and activities within a city via a globe shaped device. In the study, a method of displaying ‘seas’ in areas that are not visited by the user can be used as a means of allowing the user to investigate the contents of areas that they do not usually visit. Using such a visual representation whereby a real world feature is used to identify a specific characteristic in the area, multiple users would be able to immediately start making assumptions about what they are seeing – the sea in the real world would invoke a thought of a featureless expanse where people do not normally go.

Andrienko & Andrienko (2011) discuss a method of displaying large amounts of movement data by aggregating specific points of multiple trajectories and then using these as central points for Voronoi tessellation. Trajectories are then visualised by joining these centroids using arrows of varying style depending on the number of trajectory segments being represented. The main benefit of this study is in representing large amounts of data (in this case 6,200 trajectories) in a simple vector based visualisation, which is primarily accomplished using the clustering method of key points in trajectories.

Another technique used for visualising data is that of space-time plots and data cubes. Song & Miller (2011) define space-time plots as:

> *"a graphic representation of a variable's behaviour with respect to a given spatial aggregation level and temporal granularity"* [pp 2]

They also describe a benefit of data cubes as being:

> *"The data cube leverages space-time plots by facilitating a high degree of user-interactivity"* [pp 2].

In effect, the data cube is a structure containing multi-dimensional data that can be easily retrieved and analysed by a user. In the study presented by Song & Miller (2011), data cubes were used as a means of analysing traffic data where clear patterns could be seen between date, time and location. Space-time prisms are another visualisation that can be used to analyse potentiality of human activity within a given space (Kuijpers et al. 2010). The space-time prism visualises all possible locations a person can in be relative to their starting location, travel velocity and time taken (Kuijpers et al. 2010). The visualisation was also applied on a much larger scale and speed by Hawking (1988) to show the possibilities of causal effects within the universe based on the speed of light (as nothing can travel faster than the speed of light, anything outside of the prism could not possibly be affected by the item being observed over time). In Figure 4a an item is shown at the origin of the space-time prism. In this example, the two cones identify the possible location (x and y coordinates) of the item at a given time in the future or in the past. Note that this assumes a constant maximum velocity of movement – if the maximum velocity was greater, then the discs within the cones would increase in size quicker whereas if the maximum velocity was smaller the discs would grow in size slower. If we know the end point of the item, then the prisms can be used to identify possible locations of the item given a time (Figure 4b). Kuijpers et al. (2010) discuss that a problem with the classical space-time prisms are that they assume that the start and end points of the individual's time window are fixed and known (prism anchors). This may not be the case owing to measurement error or flexibility and so they implement a more complex prism based on networks with uncertainty and probabilities. Kobayashi et al. (2010) further the analysis of space-time prisms by developing an analytical approach that can assess error between intersections of prisms, a particularly useful tool for determining items such as probabilities of two people being in the same location at a point in time, where inherently some vagueness is present (i.e. inaccuracies in measuring time, travel speed).

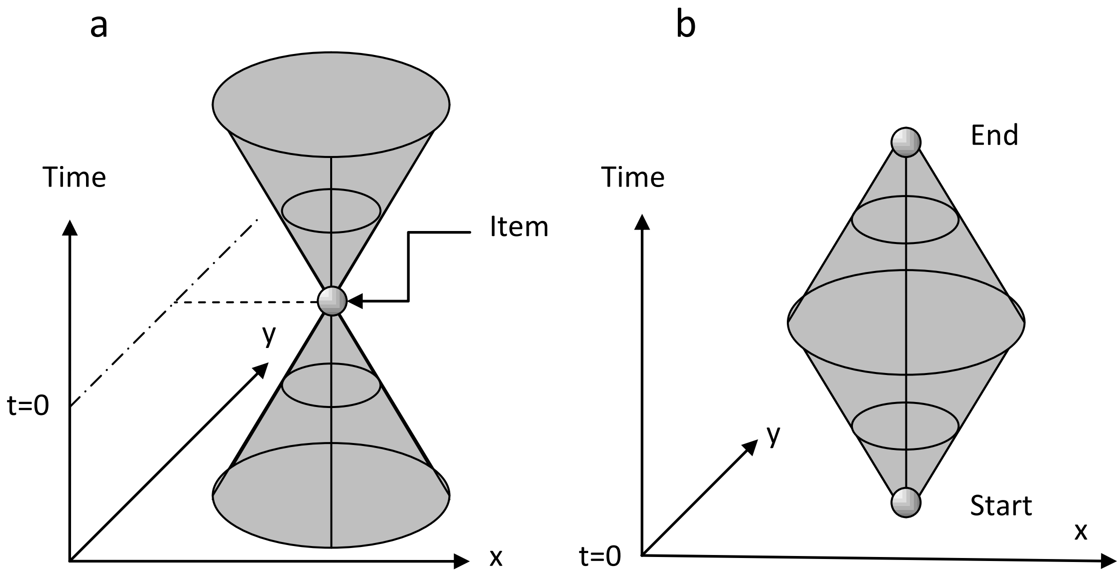


Figure 4: Space-time prism a) area of affect b) possible location between start and end points

3 Rationale

The literature review has shown that although extensive research into the psychological and physiological aspects of how distance is perceived and estimated has been conducted, little of this has addressed how the information about distances can be directly portrayed to the user. A large proportion of the literature focuses on the more natural aspects such as shadowing and binocular disparity, with some research in regard to computer science and virtual reality which in itself could be used as a platform for implementing tools to aid in the estimation of distances. Such tools could be implemented within VR or AR (Augmented Reality) as a means to provide accurate information to the user about heights and distances of objects and thus aid in orientation tasks, as well as information and analytical tasks where knowing the distance may be of value. The study by Milios et al. (2003) does however look at a tool for determining distance in the form of a laser range finder converting distances to sounds, although this study was performed in the real world and not a VE. This highlights two possibilities for portraying distance information within a VE – sonification (portraying distances as sounds) and using a virtual representation of a real world tool (laser range finder).

Additionally, most research carried out has been on shorter distances, namely within action space (up to 30m from the observer), with little research on the larger scale vista space (over 30m from the observer) as can be seen in Table 4. From the literature there appears to be little research with regard to distances within personal space (within arm's reach of the observer). This is not surprising as within this region the observer can physically touch the features to determine distance as opposed to calculation from cues in the environment.

When navigating within large scale environments, knowing distances from features such as mountains within vista space would be of great benefit as being able to accurately estimate distances is an essential part of navigation (Witmer & Kline 1998). As well as the limitation in the literature on estimations in vista space, most of the research performed is within structured environments such as buildings. As can also be seen in Table 4, a predominant number of experiments occur either in an indoor environment (in that either the VE setting or the real environment is located inside a building) or in an abstract environment (the environment used in a VE where there is no real world comparison i.e. it is totally abstracted from the real world).

When using naturally occurring cues for determining distance, it can be easily determined that many will not work over large distances. For example, over a large distance textures on features would no longer be distinguishable due to compression of the patterns. Using the size of familiar objects would also be problematic as at greater distances, the objects would either become too small to determine differences, or so small that they disappear from view altogether (the same is applicable for shadowing). We also know that many of the binocular cues such as stereopsis, accommodation and convergence become of little use over larger distances. The important thing to note here is that many of the cues that could be used at close range distance perception would not work for larger scales unless the cues were scaled themselves to take into account the effects of Emmert's law.

Author(s)	Study	Distance	Environment	Space
Allison et al. (2009)	How effective binocular depth discrimination and estimation is beyond interaction space	Up to 18m	Abstract/Indoors	Action
Clément et al. (2008)	Effects of microgravity on distance and depth perception using stereoscopic vision	-	Abstract	Action
Hendrix & Barfield (1995)	Differences between stereoscopic and monoscopic depth cues	-	Abstract	Action
Hubona et al. (1999)	How shadows effect placement and sizing of object tasks within VR	-	Abstract	Action
Interrante et al. (2008)	Looking into judgements of distance using a VE representation of a real-world location	Up to 20m	Indoors	Action
Jurgens et al. (2006)	Depth cues that can be used to improve accuracy of placing a stake in a specific location	-	Abstract	Action
Kitazaki et al. (2008)	How depth cues effect perceived lightness	8.5 – 17.5m	Abstract/Indoors	Action

Uehira et al. (2007)	Development of a display which uses motion parallax to aid in distance perception whilst in a moving car	20 – 40m	Urban	Action
Willemsen et al. (2008)	How manipulating stereo viewing conditions can effect distance perception	5 – 15m	Indoors	Action
Witmer & Kline (1998)	Investigation into the contribution of different cues for judging perceived and traversed distances within a VE	40 – 370m	Indoors/Abstract	Vista

Table 4: Spatial size of previous studies

Another note from the literature reviewed regarding distance estimations is that there appears to be minimal amounts of research on perceived heights in both VR and the real world. The study by Wan et al. (2008) showed that when estimating heights, people were more accurate when the scale used was 1:4 as opposed to 1:1. This in itself is interesting as it may indicate that a scaling of the environment itself is required for a VE to appear more realistic. In many cases, especially in a geography context, it is important to be able to estimate heights of features such as trees or buildings to be able to accurately perform certain experiments or create representations.

Ultimately it is clear that although extensive research has been conducted with regard to the psychological and physiological aspects of distance and depth estimations, there has been little with respect to large distances and the tools that could be used to aid in these. Many of the previous studies have made use of VR as a setting for the experiments conducted, but again such environments have only been used as a replacement and no tools were implemented to actually aid in the determination of distances. Therefore, could distance estimations within a VE be improved by the provision of a visual tool providing information about the layout of the environment, such as an overview map?

Research Question: Does providing an overview map inside a Virtual Environment aid in estimating observational distances?

Research has also shown that skills learnt within a VE can transfer from the virtual to real world settings. However, does providing information in an attempt to improve a task inside the VE improve general distance estimations in the same landscape when those aids have been removed?

Research Question: Does providing an overview map in one instance improve distance estimations in a second instance when the map has been removed?

Another aspect investigated in the literature review was that of navigation and landmarks. It is clear that landmarks are an important feature with regards to navigating within an environment both as a mean of indicating turning points along a route and as confirmations of whether the navigator is travelling along the correct route. In the specific case of urban environments, how we perceive features is largely dependent on how we interact with them. With the UIT, the urban environment can be split into a series of different feature types, each of which having specific characteristics in how they are used. In terms of navigational tasks we see that the category of a feature can change depending on the task undertaken, although as a general image the features would remain constant. Vinson (1999) indicates that all of Lynch's feature classifications should be included in VEs to aid in navigational tasks, but what if the environment is not urban in nature – can less well defined features take the place of the usual features being classified?

Research Question: What features in a rural Virtual Environment can take the roles of the Urban Image Theory's five feature types?

Most investigations into navigation within VEs tend to occur in environments that are contained and well structured such as rooms, buildings and urban areas. Table 5 shows a number of previous studies and whether experiments within them were conducted in virtual or real environments, and the general type of that environment. The environment type column identifies information regarding the realism of the environment (abstract/ realistic), the scale (large-scale, room, building), and general structure (maze, urban).

Study	Real/Virtual	Environment Type
Aoki et al. (2008)	Virtual	Structured –Space Station
Biggs et al. (2008)	Virtual	Game; Large Scale; Abstract
Bosco et al. (2008)	Virtual	Abstract; Room
Giudice et al . (2010)	Virtual	Building; Audio (non-visual)
Harrower & Sheesley (2007)	Virtual	Landscape; Realistic
Head & Isom (2010)	Virtual	Abstract; Maze
Janzen et al. (2001)	Virtual	Abstract; Maze
Kelly & Bischof (2008)	Virtual	Abstract; Room
Meilinger & Knauff (2008)	Real	Urban
Morganti et al. (2007)	Virtual	Realistic; Building
Mueller et al. (2008)	Virtual	Maze
Palermo et al. (2008)	Virtual	Realistic; Urban
Parush & Berman (2004)	Virtual	Room; Realistic

Pierce & Pausch (2004)	Virtual	Realistic; Large-Scale (rural & urban)
Ruddle & Péruch (2004)	Virtual	Abstract; Maze
Zhang (2008)	Virtual	Large (2km x 2km); Abstract

Table 5: Types of environments used in navigation tests

From the studies shown in Table 5, there appears to be less research into navigational practices in rural environments which are inherently less confined and structured. In fact Bidwell & Lueg (2004) state:

> *"Landmarks in a built environment, such as a university campus, which presents an image with a distinctive structure featuring places connected by paths and dominated by right angles may be qualitatively different from landmarks in natural environments."* (pp 47)

How we perceive our surroundings is also largely to do with social and dynamic aspects as opposed to the physical presence of features within it. For example, with a shopping street we are more likely to see it as a shopping street owing to people's movements and interactions and not just because there are shops located within it. Therefore, does a feature need to be a purely visible device to be used as a landmark in a navigational context, or can more contextual aspects be used by assigning meaning or attachment to that feature or area? With the abundance of geo-referenced data and software available to display it when the user is in a specific geographic location, there is already a means of providing additional information about a physical object that the user can see. Can this information be used as a means of providing additional context to features that in turn could then become landmarks?

> ***Research Question: Does providing geo-referenced contextual information aid in navigational tasks within a Virtual Environment?***

Finally, with regard to the analysis of the data collected from navigational tasks, most are based on either single or multiple explicit values that can be analysed using statistical tests. Although this is useful for determining whether the experimental aspects have had any effect, it is not as useful in determining why they may have had an effect or not. There have been studies into ways of visualising the information collected from track and event logs but they tend to represent either a single aspect at a time or focused on small numbers of users. Although space-time prisms can be used to depict information relating to movements, their implementation and use within analysis may be complicated for non-experienced users and so a more intuitive visualisation technique would be beneficial. Therefore, can an intuitive visualisation method that shows multiple pieces of information be implemented, and if so what can it tell us about the environment itself and user performance?

> *Research Question:* **Can a multi-metric analysis be performed on track log data to derive information about trial performance, how features within the environment are used in navigation tasks, and what features are present in the environment itself?**

4 Creation of the VE

Throughout this study a virtual representation of the Sorbas region in south-eastern Spain was used. The model was created using the Bionatics Blueberry3D fractal terrain generation software, with functionality added using VegaPrime C++ libraries.

Bionatics Blueberry3D is a terrain generation system that creates a 3D model based on GIS datasets such as DEMs, landcover and hydrographical/infrastructure networks. By importing these datasets and applying textures and models to the environment based on the information contained within them, realistic looking environments can be created. The display (and quality) of objects within the environment is determined dynamically using the user's viewport and the amount of stress being placed on the computer system which will be discussed in more detail in later sections.

4.1 Site Selection

The site selected for representation in this study is the Sorbas region of south-eastern Spain (a 10km x 10km region with lower left coordinates of 575000, 4100000). This site was selected for a number of reasons including intended future use of the system, data availability and the general difficulty of navigation within the region due to a lack of identifiable landmarks. With regards to future use, the real world location is frequented annually by undergraduate physical geography students from the University of Leicester as a field trip site. The overall intention was to use the environment and any tools developed as a pre-visit orientation aid for these students so that less time would be required for orientation when they arrived in the real world location. Although the region is used as a field trip site due to the geomorphologic features present, there are few remarkable terrain features or urban structures which may be used as global landmarks. Not only does this increase the difficulty of navigation within the environment, but the lack of unique features means that the environment itself should be applicable as a generic rural landscape, thus allowing transferability of skills to other rural settings. Therefore this lack of features not only makes the environment more generic but also more of a challenge with regards to navigation tasks which makes it a prime candidate for the testing of performance enhancing tools. As the primary focus of this thesis is the development of such tools and whether the implementation of these could be of benefit, the structure of the environment itself is not a key investigative factor. Obviously physical features within the environment have an impact on the tasks performed, but analysis of findings and performance is with regards to the changing circumstances of tool presentation and not particular changes within the environment itself.

As mentioned, another factor in the selection of the study region related to the availability of data. In the case of the Sorbas region used, datasets to construct the environment and photography that could be used to ground truth the visual factors produced were already available at the start of the study, which reduced time spent in the environmental data collection stage. Previous work within the department had made use of these sources of information in prior developments of similar environments. The creation of the environment used in this study drew on the work and methods undertaken to create these prior environments and as such an existing knowledge base has been extended.

4.2 Data Collection and Preparation

Before creating the environment, the data required by Blueberry3D needed to be collated and edited into suitable forms using ESRI's ArcMap GIS software. Original datasets were obtained from the University of Leicester with further data being acquired from Junta De Andalucia. Once collected, all data needed clipping to the study area using pre-existing ArcMap methods.

The data used within the Blueberry3D model can be seen in Table 6 which identifies the feature type, the feature group it belongs to, and the data type. The following sub-sections providing more detail about the manipulation methods used on each of the datasets.

Feature	Group	Type
Elevation	-	Raster
Land cover	-	Raster
Freeways	Roads	Linear
Highways	Roads	Linear
Highway junctions	Roads	Linear
Minor roads	Roads	Linear
Streets	Roads	Linear
Footpaths	Roads	Linear
Rivers	Hydrological network	Linear
Boulevards	Hydrological network	Linear
Stationary water bodies	Hydrological network	Polygon
Wire fences	Boundary	Linear

Stone walls	Boundary	Linear
High tension power lines	Power lines	Linear
Low tension power lines	Power lines	Linear
Building outlines	Buildings	Polygon
Churches	Buildings (individual)	Point
Country houses	Buildings (individual)	Point
Crosses	Buildings (individual)	Point
Swimming pools	Buildings (individual)	Point
Service stations	Buildings (individual)	Point

Table 6: Data used in Blueberry3D model

4.2.1 DEM

Initial elevation data were in the form of contour lines which needed to be converted to a raster representation for import into Blueberry3D. Using an inbuilt ArcMap tool, the contours were interpolated into a 5m resolution raster grid with each cell representing the terrain elevation of that point. This raster was then exported as a geotiff image ready for import into Blueberry3D (Figure 5).

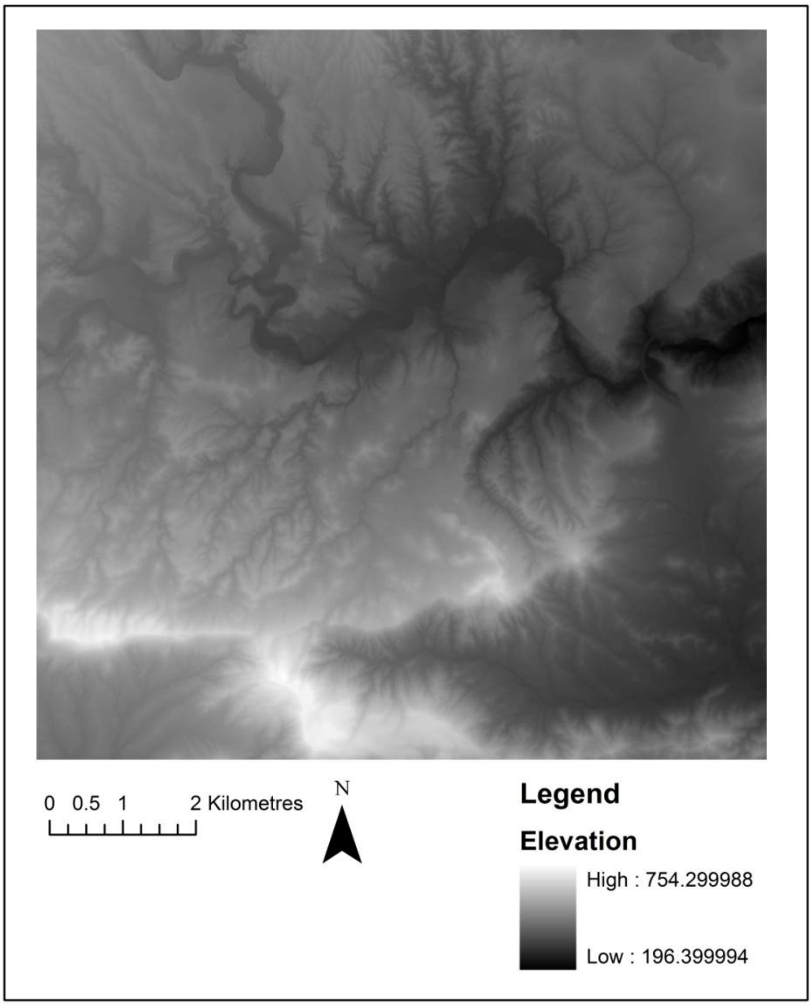


Figure 5: DEM

4.2.2 Road and Hydrographic Network

Road datasets were available in the form of three ArcMap vector shapefiles (primary roads, streets, and minor roads and footpaths) with each containing further sub sets of road classifications identified with database codes on each feature. As Blueberry3D requires individual shapefiles for each road classification, the different road types needed to be extracted based on type codes within each record using SQL commands. Upon completion, the different road types extracted were freeways, main roads, primary junctions, minor roads, streets and footpaths. To allow for better display in the model, the road vector lines were split whenever they crossed a 5m elevation contour to ensure that when heights were assigned to line vertices within Blueberry3D, a suitable number of vertices existed to reduce the amount of ‘cutting through terrain’ artefacts (where the roads create a channel through a hill). Additionally, some road shapes comprised of three lines (one for each carriageway and a central line) which needed aggregating so that only the line representing the centre of the road remained.

Additional datasets representing the hydrographical network of the region were present which held information about rivers, streams, boulevards and stationary water bodies. As the study region is predominantly arid, it was decided to exclude the streams from the model as these would be dry for the majority of the year. Although river datasets were in the form of polylines, the boulevard dataset was in the form of polygons. For the creation of flowing rivers, Blueberry3D requires polyline data and so the polygon boulevard data needed to be converted. By converting the features to lines which provided a linear outline of the boulevards, and then extracting the centre line of these boundaries, a single line representing the boulevard was generated.

The complete road and hydrography network can used can be seen in Figure 6.

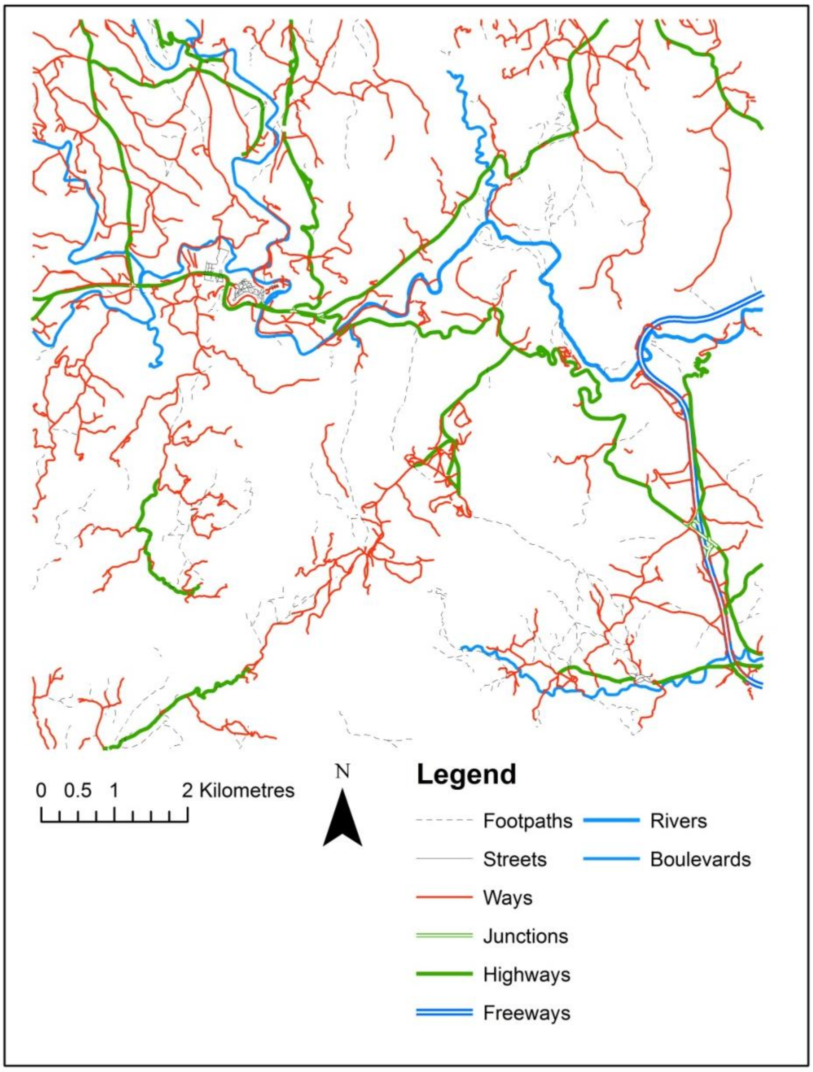


Figure 6: Road and river network

4.2.3 Power Lines and Boundary Features

Power line data existed in a single polyline Shapefile which also stored a numerical field identifying the tension of the line. This field was used to split the data into high tension (132 in the tension field) and low tension (20 or 25 in the tension field) so that the two types could have a different representation in Blueberry3D. The boundary data consisted of linear features representing the location of wire fences and stone walls, which were in a form that required no additional manipulation. Figure 7 shows the power line and boundary feature data used.

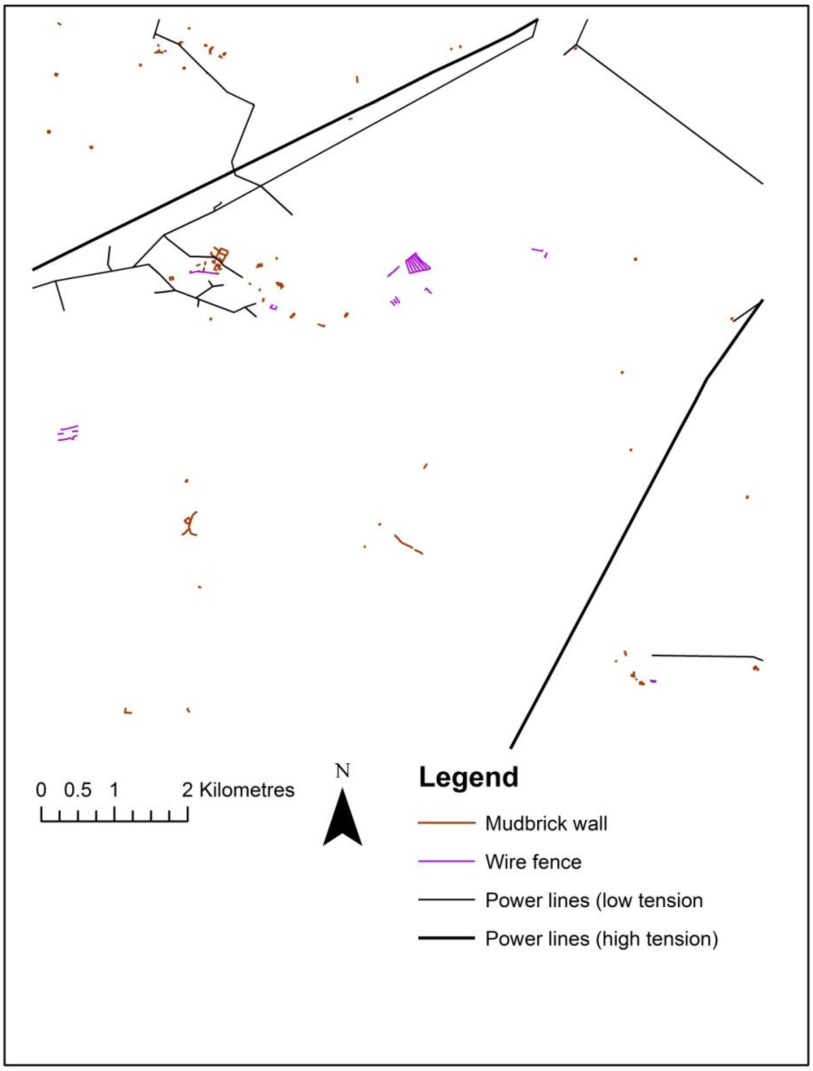


Figure 7: Power lines and boundary features

4.2.4 Buildings

Data about buildings were present in both polygonal and point data format. The polygonal dataset represented the outlines of buildings within the region, and the point dataset indicated the location of specific buildings or structures such as churches, country houses and crosses. This combined point dataset was separated into a number of point datasets which each contained different structure types (i.e. one dataset for churches, one for country houses etc.). As these point datasets identified the location of models within the environment, point in polygon methods were used to remove their outlines from the polygon dataset. These methods were implemented using inbuilt functionality of ArcMap whereby all polygons of the building outline dataset that contained one of the points from the individual building datasets were removed. The remaining polygons were then split into 4 datasets based on their ID (modal division of the ID field) to allow for four different building visualisations within the environment. The polygon and point features used can be seen in Figure 8.

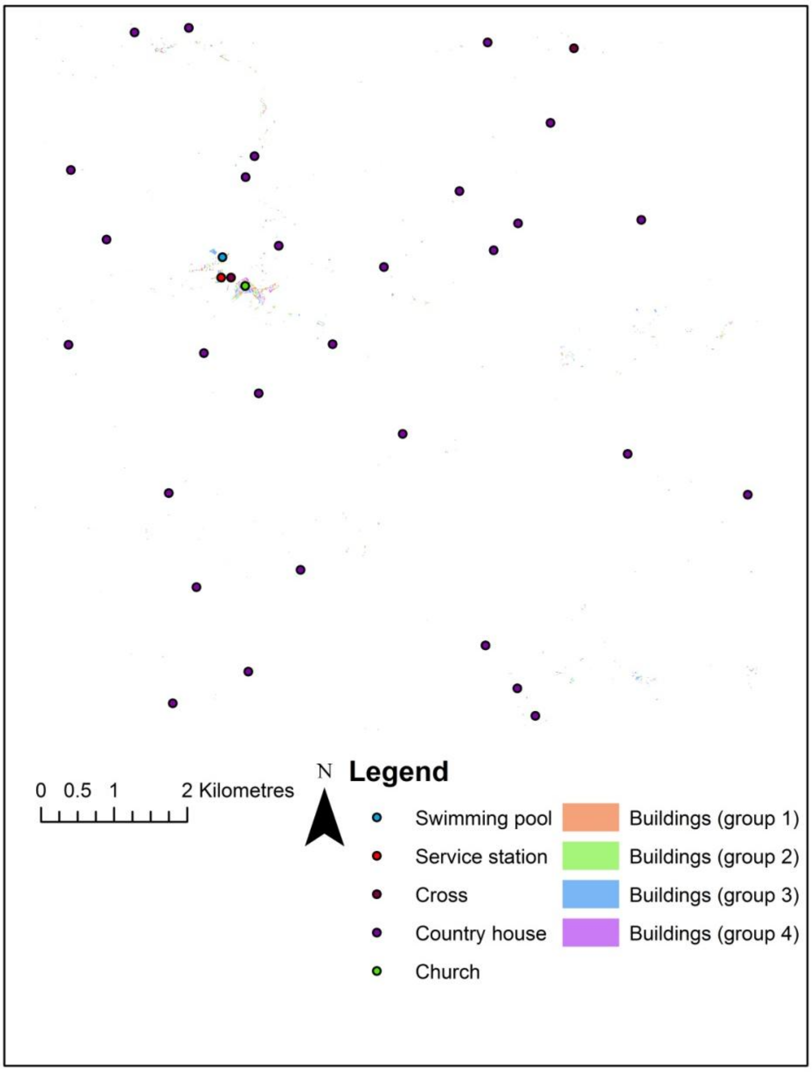


Figure 8: Buildings

4.2.5 Land Cover

The land cover dataset was in vector polygon form, with each shape representing a region and having a land cover assigned to it. This dataset was first converted to a raster using ArcMap inbuilt tools before exporting as a tiff file with each land use having a unique colour assigned to it (Figure 9). Within Blueberry3D, the colours used in the image are used to identify which land cover is to be used at that location. It should be noted that a different land use dataset was used for the first experiment (distance estimations) as newer and more detailed data was made available before the second experiments took place, although the same classifications within Blueberry3D were used.

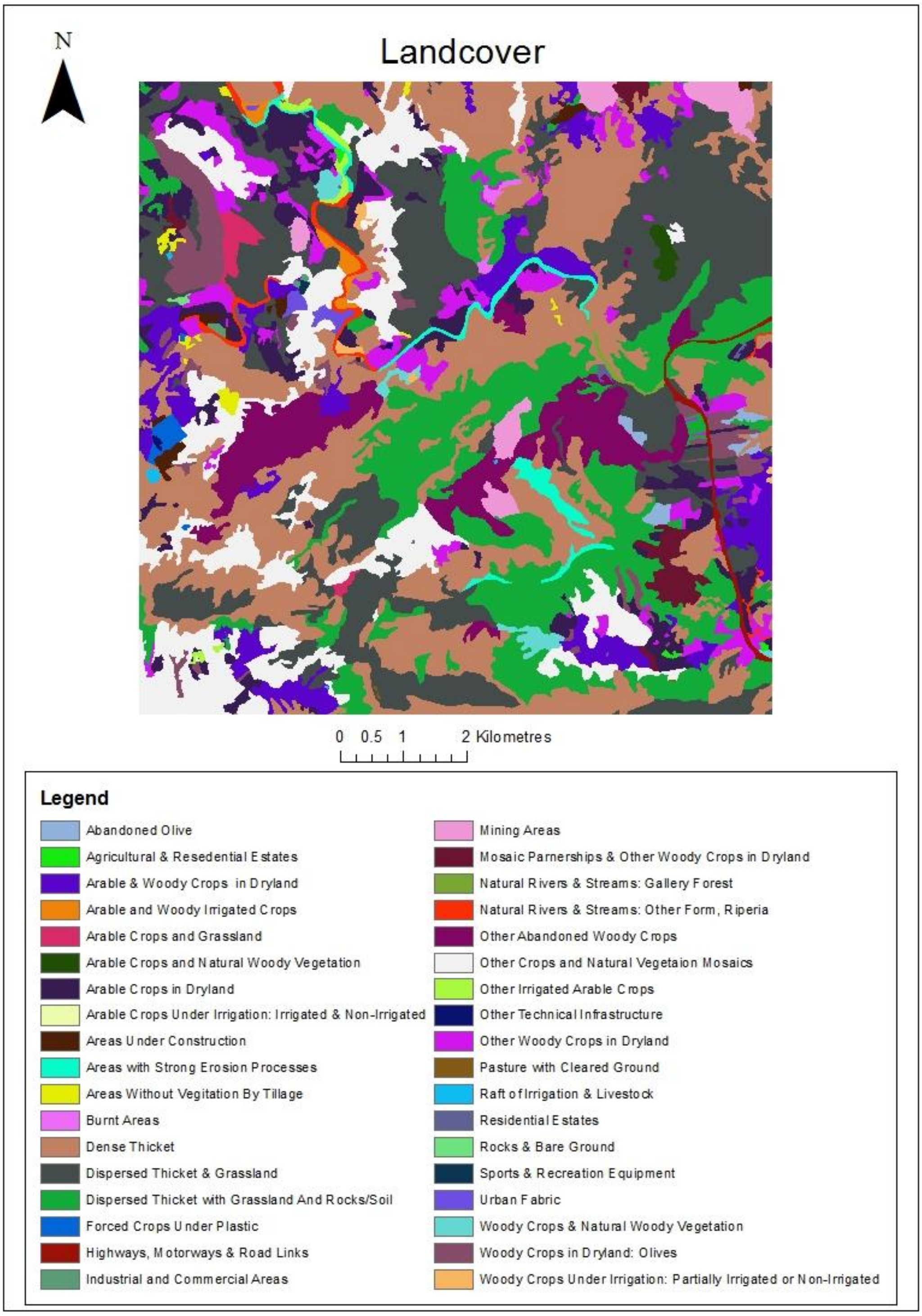


Figure 9: Land cover

4.3 Modelling

The visual representations of features within Blueberry3D generally either required flat textures from image files (photographic or user generated imagery placed onto surfaces) or 3D models. For the 3D models, Presagis Creator was used and the models exported as open flight files (.flt). One model was created for each structure type (with the exception of the service station where an existing model was used) based on images viewed from the internet and photography from previous field visits. Examples of the models generated in Creator can be seen in Figure 10.

a b

c d

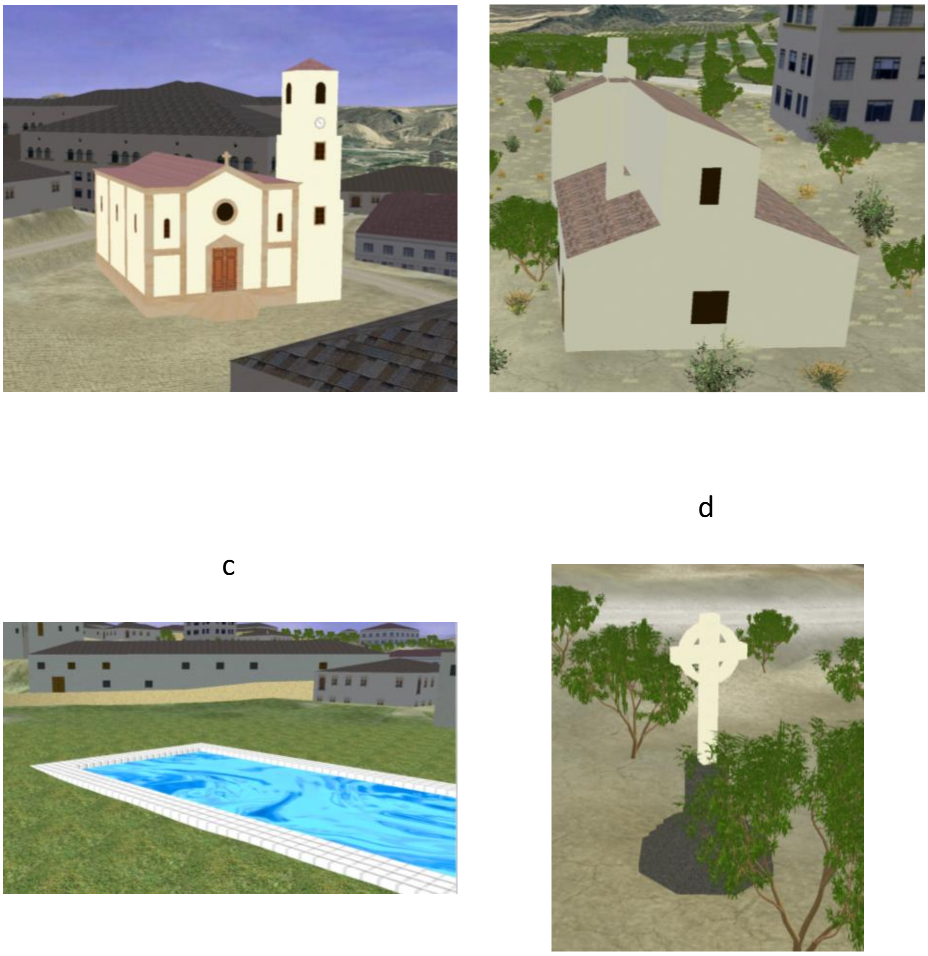


Figure 10: 3D models created a) church, b) country house, c) swimming pool, d) cross

Linear features in the environment such as roads, rivers and fences required the creation of textures for placing on them. In some instances suitable textures were already present such as the highways, but in other cases they required creation using

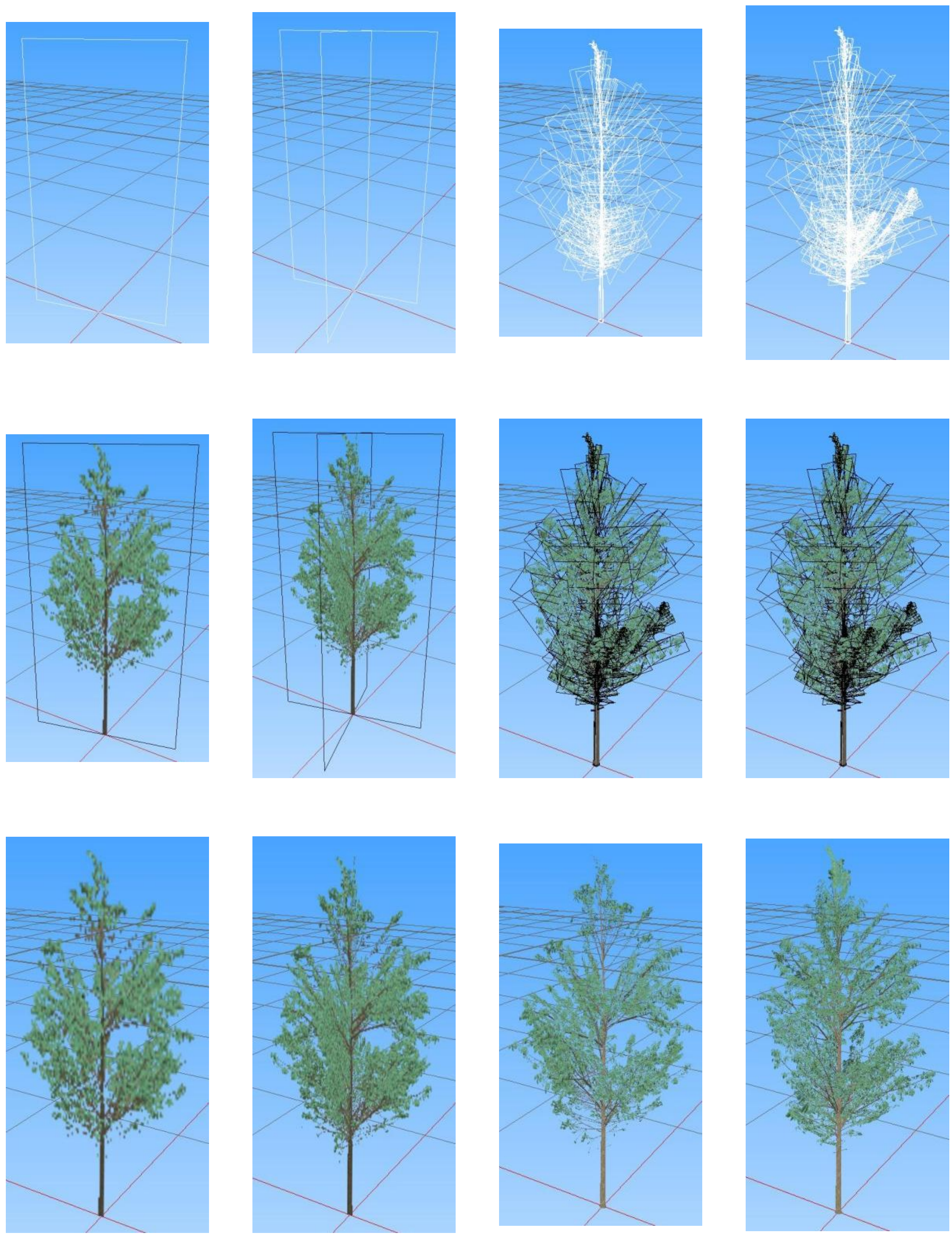
Adobe Photoshop. Within Photoshop, various techniques and filters were used to create water textures for use on the rivers, road surface textures for freeways, minor roads, streets and footpaths, stone texture for stone walls, and a partially transparent wire texture for wire fences. Additional textures were also created for the ground surfaces if not already present, as well as using photographs to create textures for the extruded polygon buildings.

4.4 Blueberry3D Environment

As mentioned, Blueberry3D generates a realistic looking environment by collating various datasets into a geographic database and then determining visualisations based on this data and visual item representations provided. As such, the database was set up using the same projection system as the GIS data that were imported using inbuilt tiling procedures (splitting a large region into multiple smaller regions). By tiling large datasets the system can determine which tile to display at any given time and so reduce computational costs. To further increase performance, Blueberry3D uses 'Levels Of Detail' (LOD) to reduce the amount of polygons used for object display depending on their distance from the observer. Rather than using physical distances, Blueberry3D uses octaves to determine LOD, which represent the size of triangles drawn on the terrain surface. For example, at a 'distance' of 0 octaves, the triangles are closer to the observer and so need to be smaller to provide a realistic surface. Therefore the ground surface triangles drawn are with a size of between 1 and 2 metres. At an octave value of 1, the triangles would be between 2 and 4 metres in size and at octave 2 they would be between 4 and 8 metres. The size of the ground surface triangles can be determined using $size = 2^{octave-1}\ to\ 2^{octave}$. The main benefit of using octaves as opposed to physical distances is that if the system is under high strain then the octaves can be modified accordingly so that larger octaves (bigger ground triangles) are used even when the observer is close. This in turn reduces the amount of triangles being rendered and so increases performance.

Although octaves are related to the ground surface, their values are used to identify the LOD boundaries for all item representations. Complex features such as trees generally have four LODs, each with a different model associated with it. For the most coarse representation, a simple billboard representation is used (an image on a single plane which rotates depending on its location according to the observer), followed by a crossboard (2 planes forming a cross shape), followed by a low polygon model, and then a high polygon which is shown in the images of Figure 11.

Billboard	Crossboard	Low Polygon	High Polygon



Triangles: 2	Triangles: 4	Triangles: 521	Triangles: 1108
Vertices: 4	Vertices: 8	Vertices: 1510	Vertices: 3564

Figure 11: LOD modelling

As can be seen by the images and triangle/vertices count in Figure 11, it is important to define changeover values where the increase in workload is required. For example, you would not use the high polygon model at distances of greater than two or three metres as the massive increase in triangles is only of benefit if you are at a distance where the individual features (i.e. leaves) would be distinguishable from each other. Similarly, you would not have the billboard model in such close proximity as there would be no 3D effect and the representation would look very poor. Again however, as the representations are controlled by octaves and not explicit distances, the distance at which the LODs swap may change depending on the strain on the system. If there are a large amount of trees being rendered, then the system may postpone the change to high polygon models until the observer is much closer to the particular feature. Dockerty et al. (2005) note that 2D representations of vegetation are more commonly used in virtual representations of rural landscapes due to their numerous occurrences. Bluberry3D allows the use of both the 3D and 2D representation at different scales and so the problem with performance for showing more realistic models is reduced. Appleton et al. (2002) state that users have to make a trade-off between levels of detail and interactivity based on resources that are currently available. Even though this is still the case today (their study was in 2001-2002) resources have vastly increased with far superior processing power and memory availability, along with the software available (Dolman et al. 2001). Software such as Blueberry3D aim to make use of these greater resources by creating more realistic representations which can be interacted with, but there are still severe restrictions with regards to what level of realism can be presented and costs in term of money and time to learn the software (Appleton & Lovett 2003).

4.5 Land Cover

Different land cover types within Blueberry3D are created by combining different item representations (trees, bushes, rocks etc.) with surface covers (grass, sand, soil etc.) in varying levels and densities. For example, a thicket land cover was created by combining bare earth and grass surface covers with various shrub representations. To create a dense thicket layer for another land cover classification, the same combination of item representations and ground covers were used, but the density of the shrubbery and amount of grass cover were increased (Figure 12). Once created, the land cover visualisations for all land cover types were mapped to the RGB colour values of the tiled land cover classification image created using ArcMap (as seen in Figure 9).

Dense thicket terrain classification	**Dispersed Thicket & Grassland terrain classification**
	

Figure 12: Terrain classifications

After the first experiment was conducted, visual updates were made to the environment. As mentioned earlier new land cover data became available for use but also modifications were made to some of the existing Blueberry3D terrain cover definitions to make the environment appear more realistic as well as improve performance. For example the grass billboard images were updated to change the colouring from green to a browner shade, as well as using the manual override on item representation's BLOD (Blueberry Level Of Detail) octaves to make the more detailed models start to appear closer to the viewer. It was felt important to update these aspects as the second experiment would require movement through the environment. This would mean that additional processing strain would be placed on the system due to the constant updating of features that are to be drawn on the screen as the user moves locations. This extra processing would slow down the rate at which frames are drawn on screen giving a 'jumpy' appearance to any movement and rotation, and so the amount of processing needed to be lowered by reducing the number of polygons present in any one view.

4.6 Vega Prime

To display the Blueberry3D model in stereo vision and allow user interaction, Vega Prime C++ libraries were used. Alongside the libraries, a GUI (Graphical User Interface) named LynX was available that allowed for a more simple yet restricted interface to allow for the implementation of basic extended functionality. Once basic settings were applied for the display and interaction, more specialised functionality was coded using C++.

A key benefit of the VegaPrime system is that the environment can be displayed in stereo vision. This is accomplished by rendering the environment twice from slightly different viewpoints and then the two images projected onto the same surface. In the system used for this investigation the environment was projected through two projectors each using a different light polarisation. The projected images appeared on a screen made of specialised material to maintain polarisation and the user wore filter glasses to only allow specific polarisation to reach each eye. A benefit of this system was that the equipment remained portable and relatively compact, but the room needed to remain dark as large amounts of light from the projectors were lost in the polarisation process. Additionally the scene needed to be rendered twice on the computer system and so used twice the amount of memory and processing power as opposed to a flat projection. Owing to the amount of processing required, the realism of the model needed to be balanced with performance as a large number of polygons would in turn reduce the frame rate making the movements jerky.

In the experiments, VegaPrime was also used to add the custom functionality not existing in current libraries such as the display of information on screen and logging of user movements. In both of these cases, although C++ was used to code the functionality, a large amount of information (such as the user’s location) was directly acquired from the system which was stored using the VegaPrime data structures. These advanced features will be discussed in detail in the relevant chapters.

One problem with the system that led to reduced realism was the implementation of shadows. Although the VegaPrime system does allow for the rendering of shadows, when implemented the shadows produced were unrealistic in that they flickered and extended incorrect distances. Therefore to reduce computational costs the shadows were removed from the system.

4.7 Research Laboratory Setup

The room used as the experimental laboratory was a specifically designed venue for VR applications, having darkened walls, no windows and variable brightness lighting. The environment was projected onto portable screen measuring 160 cm by 130 cm which was coated in a special material to maintain light polarization. The projection itself was made through two portable projectors each filtered to a particular light polarization to create the stereo effect, and covered the entire portable screen. Users sat behind the projectors (maintaining a full field of view) with the screen placed three metres away. This ensured that no shadows created by the user appeared on the screen. Figure 13 details the layout of the laboratory.

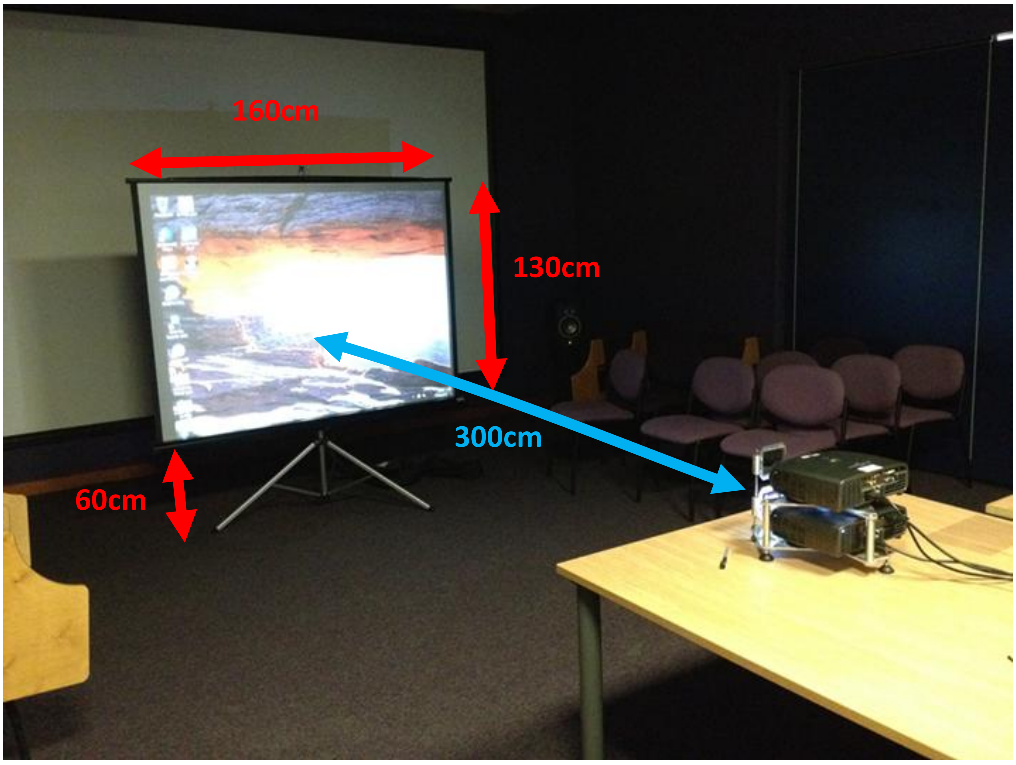


Figure 13: Layout of research laboratory

4.8 Participant Selection

Selection of test participants for the all investigations in this study originally included criteria to cover a wide base of personal and background criteria. After careful consideration it was decided to reduce this population to focus on geography students and staff – the ultimate end users of any practical tools developed from this research. The reasoning behind this alteration was due to a difficulty in determining a difference in spatial literacy between test subjects which could have an impact on performance. By restricting the population to people from a primarily geographic working background, the variation of spatial literacy between subjects is reduced as such people will have had a high exposure to geographic and spatial tasks.

In addition, time constraints and access to individuals was also a consideration in reducing the overall population covered by the investigation. No prior experience regarding the use of VR systems was required, and it was decided that no restriction should be placed on whether the test subject had been to the site previously before (either in a virtual or real sense). The reason for not placing this restriction was so that qualitative feedback could be collected regarding prior experience of the environment. Users were not explicitly asked in any experiments if previous knowledge was used. Instead the obtaining of information regarding this was through the in-trial verbalisation of thoughts. In addition, due to the large size of the environment used and the generic nature of the landscape, it was deemed unlikely that any prior experience in the environment without a focus on navigation and orientation would have a large impact on the overall performance. If however there was an impact, it was felt important to capture this aspect as a factor in the overall spatial reasoning of the tasks as this would indeed highlight the transferability of skills gained from one representation of the environment to another. This would also be an important aspect regarding the tools implemented as it is likely that the identification of the same feature would predominantly be determined through the tools rather than the environment (i.e. the naming of locations in the info-marks presented in Chapter 6).

5 Part I: Observational Distance Estimation in VR

Does providing an overview map inside a Virtual Environment aid in user estimation of observational distances?

- ***Does providing the ability to zoom in on the map affect the estimations?***

Does providing an overview map in one instance improve user estimations of distance in a second instance when the map has been removed?

5.1 Introduction

The first investigation is with regards to distance estimations in VEs. Previous research has shown that distance estimations within VEs are often inaccurate (Bodenheimer et al. 2007, Clément et al. 2008, Interrante et al. 2008, Willemsen & Gooch 2002, Willemsen et al. 2008), many noting that this could be a result of a lack of distance cues which are more prevalent in real world environments. It appears however that although this problem has been identified little has been done to counteract the inherent inaccuracy, with many suggestions simply be to make environments more realistic and to add more cues. In addition, the literature review has shown that a large portion of research focuses on the shorter distance estimations with few investigating longer distances such as those present in outdoor rural environments.

This chapter explores:

- A possible solution to the problems arising in distance estimations by providing direct feedback about the environment to the user via an overview map which has surprisingly not yet been investigated.
    - A geographer's intuition suggests that the provision of such a map would increase accuracy due to the presentation of information about the scale and location of the features in the environment.
- The benefits of allowing dynamic 'zooming' of the map.
    - By allowing the user to alter the scale of the map they can see in more detail features in their surrounding area and so may be able to more accurately assess the distances to them (via a dynamic scale bar). Although seeming fairly obvious that this should be the case, as of yet it has not (to the best of the author's knowledge) been tested.
- How the presentation of a map in one instance affects the accuracy in subsequent estimations when the map has been removed.
    - Due to the evidence of learning between tasks it is a probable outcome that by better learning distances in one task (where a map is present if this in fact increases accuracy) following tasks should also be more accurate even if the map has been removed.

This chapter details the methods used to generate the maps and to show them to the user. The quantitative and qualitative methods of analysis are also outlined and the results for this set of experiments discussed.

5.2 Methodology

The main extension of the blueberry environment for this experiment was the inclusion of a dynamically scalable overview map on the display. This map showed the user's location as well as the location of the marker to which the distance is to be estimated, with features in the environment (roads, rivers etc.) also being displayed.

Analysis was performed via verbal reporting of estimated distances between the user's location in the environment and an artificial marker inserted in the field of view. It must be noted that previous studies have found that verbal feedback may not be the most accurate method of recording such estimations (Interrante et al. 2006) as this method can introduce bias and noise (Thompson et al. 2004). Philbeck & Loomis (1997) on the other hand found no differences between verbal and walking estimations. For this study however, using a 'blind walking' method (as used in Ahmed et al. (2010) and Phillips et al. (2010) – a method where the user is asked to blindly walk to where they believe the object is located) would not be practical owing to a lack of specialised equipment and the large distances involved. In fact, Ahmed et al. (2010) note that past a distance of 66 feet the error introduced by the blind walking method is linearly proportional to the distance meaning that a large error could be expected from the distances used in this study. Statistical analysis was performed on the values given for the distance estimations, with further analysis performed on qualitative data collected both during the trials and from post-trial interviews. As well as the initial estimation, which tests whether the display of the map has an effect on the estimations, a second estimation was also obtained from each user where in all cases no map was displayed. This test is to identify whether the presentation of a map in the first estimation has an effect on an estimation where the map has been removed.

5.2.1 Overview Map

For the investigation, an overview map was needed to be displayed on the screen and the motion model customised to ensure that users could not move from their location, as well as to allow for interaction for scaling the map. The map display and interaction were implemented using the VegaPrime libraries and C++. To implement the overview map three main stages were undertaken:

1. Creation of the map texture file using GIS software
2. Adding of the map to the environment display with dynamic scaling functionality
3. Adding of a dynamic scale bar

5.2.1.1 Map Creation

The implementation of the on screen map was performed by using a map image as a texture on a geometric primitive (a square programmatically drawn on the display). For the textures two maps were created for display at high and low scales, with the first being displayed up to 1:10,000 scale (1:x where x >= 10,000) and the second being used from 1:10,000 scale (1:y where y < 10,000) at higher resolutions. There were two main reasons for using two maps rather than one:

- At higher scale values (> 1:10,000) only more predominant features are displayed to ensure that the map itself does not become too cluttered and thus would be difficult to read. More detail is shown in the lower scale map (< 1:10,000) where fewer features are visible.
- The resolution of features needs to be modified as people zoom in on the map so that they do not appear too large and pixelated.

The maps were created in ArcMap using the same data as used to construct the Blueberry3D model, meaning that any information shown on the map would directly relate to features appearing in the environment. At the lower resolution map scale (> 1:10,000), the features displayed were:

- Roads
    - Freeways
    - Highways
- Rivers/boulevards

On the higher resolution map (< 1:10,000) additional road features were displayed so that the objects on the map were:

- Roads
    - *Freeways*
    - *Highways*
    - **Minor roads**
    - **Streets**
    - **Footpaths**
- *Rivers/boulevards*

Other features such as buildings, electricity pylons and contour data were not shown on the map owing to the relatively small map window size and to reduce cluttering. As the predominant features within the environment that could be seen from a distance were the road and river networks, it was felt that these were the most important to display. Although display of elevation and slope data would also have been beneficial, the belief was that including these in the visualisation would result in the display of too much information on the map and make other features more difficult to identify. Buildings were left out of the map display as they were not frequently seen in the environment unless the observer was close to an urban area or near a singular building located in the rural area which would then be fairly difficult to identify on the map owing to the map size and relative size of the building.

Symbology for the features was based on standard display customs applied in UK road atlas maps (i.e. freeways (motorways) being shown as a single blue line with a white line running down the centre, main roads as green, and rivers as a light blue continuous line) as to provide a more familiar structure to the map for the users. This meant that a legend could be initially left out of the map display to reduce the amount of information being displayed, although the user could ask at any point what the map features represented. If the system was to be used when a person was not present to be asked, then a legend would need to be displayed, especially if the person undertaking the experiment was not accustomed to the conventions used.

Once the generation of the maps in ArcMap had been completed, they were exported as JPEG images. The maps were both expanded to full extent before being exported, and the whitespace cropped out using photo editing software. The map used for the lower scale display (< 1:10,000) was exported using a higher DPI (Dots Per Inch) value to ensure that lines remained smooth when the map was zoomed in which although improving detail vastly increased the file size of the texture. As both maps were exported at the same scale, as long as this original scale was known the correct scale of the current display could be calculated at any time on either map.

5.2.1.2 Map Display

As mentioned earlier, the maps generated were to be implemented in the environment as textures on a geometric primitive. Owing to the nature of how textures are placed on such objects, the size and position of the texture could be directly manipulated by the system to ensure that the map was always centred on the user's location in the environment, and that they could zoom in and out of the map.

The primitive used to show the map texture was a square shape fixed to a set size and position on the window. It was important that the primitive appeared square on all displays (no matter the dimensions of the screen) so that the scale bar could be directly translated to the map and the features displayed were a true representation. This was done by calculating ratios of the screen dimensions through programmatic queries of the system and then drawing the primitive based on these values[6]. The two map textures used were loaded into the system at startup and stored in memory ready for swap over when the predefined scale threshold was passed. The positioning of the texture on the primitive surface was determined by the user's current location in the environment so that the user position was always centred on the primitive display (a coordinate value of 0.5, 0.5 on a primitive who's axes are 0 to 1 in both the x and y directions). In the case that the texture filled the primitive exactly and the map was 10,000m by 10,000m, then for the user to be centred on the map (their environment location of x=5,000, y=5,000), the texture's origin (bottom left corner) would be drawn at 0, 0 on the primitive. If however they were located in the south-western region of the environment (i.e. x=1,000 y=1,000) then the texture would need to be shifted

[6] The default for VegaPrime is to use decimal numbers between -1 and +1 as a coordinate system of the display as opposed to actual pixel values

accordingly so that their location was still centred on the primitive. This can be calculated by taking a ratio of the environment size and primitive axis size and then calculating the map position based on this. The equation for calculating the origin coordinates is:

$$O_{(x,y)} = P_{(x,y)} - \left(\frac{E_{(x,y)}}{R_{(w,h)}}\right)$$

5.1

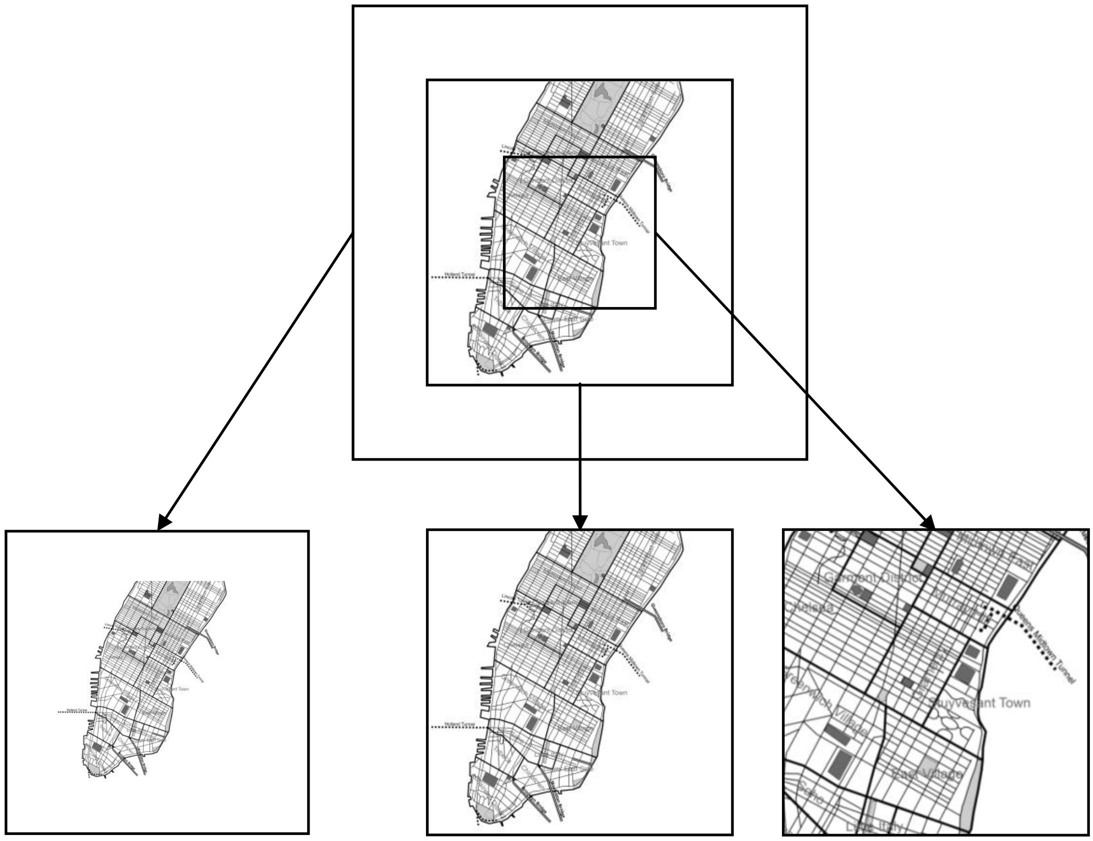
Where $O_{(x,y)}$ is the map's origin on the primitive, $P_{(x,y)}$ is the user location on the primitive (in this case x=0.5, y=0.5), $E_{(x,y)}$ is the users location in the environment, and $R_{(w,h)}$ is the overall dimension of the environment. For example, using the location of x=1,000 and y=1,000 in a 10,000m^2 environment and a primitive (and map) axis of 0-1 along both axes, we can derive that on the map itself the user is located at coordinates x=0.1 and y=0.1. As the map needs to be centred on the user's location, we shift the map texture on the primitive. As mentioned earlier, if the map filled the primitive then we would draw its origin at 0, 0 on the primitive and thus the centre of the map and primitive would be at coordinates x=0.5 and y=0.5. As the user is no longer located at the centre of the map, we cannot say that the primitive centre is equal to the map centre. Instead we need to shift the map so that the map coordinates of x=0.1 and y=0.1 are located at x=0.5 and y=0.5 on the primitive. Therefore, we set the origin of the texture to be at x=0.4 and y=0.4 on the primitive.

If we apply the earlier equation to this example, we get the following:

$$O_{(x,y)} = P_{(0.5,0.5)} - \left(\frac{E_{(1000,1000)}}{R_{(10000,10000)}}\right) = O_{(0.4,0.4)}$$

5.2

To achieve the zooming effect, methods to alter how much of the map image texture was used to fill the geometric primitive were implemented. In the simplest case where the whole map image was used to texture the primitive, upon declaring the texturing element within the programming code coordinates of the origin (lower left) would be declared as (0, 0) and the upper right as (1, 1). To zoom in and out on the map, less of the texture used to fill the polygon would be required. For example, if we wanted to double the size of features in the map we would tell the system to use only a quarter of the texture to fill the primitive (Figure 14). Any amount of texture that is outside of the primitive boundaries would not be displayed. Similarly to halve the size of the features the system would display the whole as being a texture on only a quarter of the primitive. Counter-intuitively, in the VegaPrime system the texture coordinates are not the location on the primitive that the texture is drawn on, but in effect the coordinates of the texture mapped to the upper and lower coordinates of the primitive. For the zoomed out example in Figure 14 -0.25,-0.25 maps to 0,0 on the primitive and 1.25,1.25 maps to 1,1). The location of 0,0 on the texture is at 0.25,0.25 on the primitive, and 1,1 is at 0.75,0.75.



Texture Coords: (-0.25, -0.25) to (1.25, 1.25)

Texture Coords: (0, 0) to (1, 1)

Texture Coords: (0.25, 0.25) to (0.75, 0.75)

Figure 14: Texture scaling

Assuming that it is always the centre of the texture (x=0.5, y=0.5) that is being zoomed in on and that the full size texture fills the primitive, the following equations to calculate the upper and lower texture coordinates being mapped onto the primitive can be used:

$$LT_{(x,y)} = 0.5 - \frac{1/_z}{2}$$

$$UT_{(x,y)} = 0.5 + \frac{1/_z}{2}$$

5.3

Where $LT_{(x,y)}$ and $UT_{(x,y)}$ are the lower left and upper right coordinates respectively, and z is the zoom level where $z = 1$ is no zoom, $z = 2$ is a zoom in doubling the size of features, and $z = 0.5$ is a zoom out halving the size of features. This zoom value (or scale factor) can be derived by the ratio of the new scale required against the original scale of the map:

$$z = \frac{new\ scale}{original\ scale} \tag{5.4}$$

If it is the case that the scaling is not being performed with relation to the centre of the texture then the texture needs to be translated so that the scaling origin is at the centre of the primitive and then the scaling performed. Finally the texture would need translating back so that the centre of scaling returns to the original position on the primitive.

5.2.1.3 Scale Bar Creation

A more complex aspect of displaying the map on screen was the development of the scale bar. Owing to the dynamic nature of the zoom feature and the fact that screens of varying size and resolution can be used to display the system it was important that the size and measurements contained within the scale bar were dynamically altered by the system. For this dynamic transformation, information about the screen resolution and width was needed. The user upon system initialisation entered the width of the display as the computer settings are not the only factor that alters this value (i.e. a projector can be made to display a screen at any size within its focal range) but the resolution can be directly accessed from the computer within the program. Once the window containing the environment is created on the screen, information about its size needs to be determined which will be returned in pixels. Once obtained, these values can be used to ensure that the map always stays the same size on the screen no matter what the window size is. This is done by determining how big one pixel is with regards to the window size of -1 to 1. Firstly, it is required to derive the centre of the window in terms of pixels on the screen. The centre is located halfway along the window width so this can be derived using:

$$C_{(px)} = \frac{pw}{2}$$

5.5

where $C_{(px)}$ is the centre of the window in pixels on the screen, and pw is the width of the window in pixels. Once this value is known, the size of a pixel with respect to the window coordinates of -1 to 1 can be derived using:

$$S_{(x)} = \frac{1}{C_{(px)}}$$

5.6

As it is already known that from the centre of the window to the edge is a value of 1. The same calculations are then applied to the height of the window to identify the size of a pixel on the window against the y axis. For example, if the screen has a resolution of 1920x1200 pixels, the window is displayed on the screen from (100, 50) to (1200, 1000) and we want to show the map at a fixed size of 200x200 pixels we get the following:

$$pw = 1200 - 100 = 1100$$

$$C_{(px)} = \frac{pw}{2} = \frac{1100}{2} = 550$$

$$S_{(x)} = \frac{1}{C_{(px)}} = \frac{1}{550} = 0.0018$$

$$ph = 1000 - 50 = 950$$

$$C_{(py)} = \frac{ph}{2} = \frac{950}{2} = 425$$

$$S_{(y)} = \frac{1}{C_{(py)}} = \frac{1}{425} = 0.0024$$

$$W_x = 200 \times 0.0018 = 0.36$$

$$W_y = 200 \times 0.0024 = 0.48$$

5.7

Where W is the determined width in x and y directions of the map in the VegaPrime screen coordinate system. Therefore, to draw the map so that its top right corner is at the extreme top right corner of the window (1, 1), we would define its coordinates as:

$$L_{(x,y)} = 1 - W_{(x,y)}$$

5.8

giving values for coordinates of (0.64, 0.52) to (1, 1).

To display the scale bar, it is first required to determine the original size of the map based on the screen it was created on. This is determined using the following equation:

$$M_w = \frac{px}{sr/_{sw}}$$

5.9

Where M_w is the original width of the map in mm, px is the original width of the map in pixels, sr is the screen resolution the map was created on, and sw is the screen width in mm that the map was created on. This then needs to be compared with the size that the map would appear on the current screen and the ratio applied to the scale of the map.

Calculations then need to be performed to identify the difference in size that the map would appear on the current screen to that on the original. For example, if the map is to be projected on a wall at a width of 1m but was originally a size of 20cm, the scale information would need to be modified. In the original map of scale 1:1000, 1cm on the map would be 10m in real life, whereas now on the screen the same distance of 10m would now be displayed using a line of 5cm meaning that the scale is now 1:200. If it is known that the map on the current screen would appear 1m across, then the ratio of new and original size is taken (1m and 20cm) to identify that the map is 5x bigger and therefore the scale needs to also be modified (1:1000 becomes 1:200).

Once the scale ratio has been determined, the previous calculations can be used to determine how large a specified distance on the window is. As the window coordinates (-1 to 1) can be calculated based on the screen size in terms of pixels and mm, and how big one pixel is based on the scale (if 1px measures 0.01 cm on screen, then at a scale of 1:1000 10m can be drawn as a line of 1cm which in turn is 100 pixels). Additionally, as the scale factor value used to zoom into the map is known, this can also be used to alter the size of the scale bar. Figure 15 shows the same map location with the scaling technique implemented.

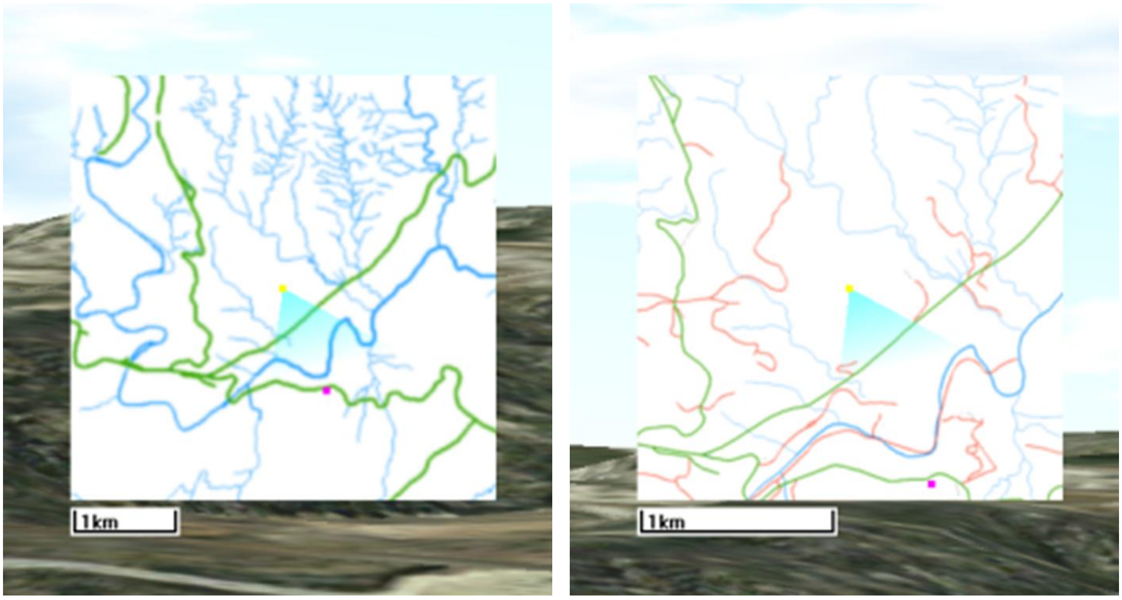


Figure 15: Map scaling

5.2.2 User Interaction

Interaction with the system was split into two forms: test subject interaction and researcher interaction. Test subject interaction revolved around the movement of the viewport and map interaction by the test subject via a computer mouse. During trials, users were restricted in movement terms in that they could not move from the viewing location – they could only look around. In addition to restricting movement, for test subjects given a scalable map interface, they could control the zoom level by clicking the mouse buttons. The system was developed so that the map could be zoomed in on using the left mouse button, and zoomed out using the right. Zooming control was also via the keyboards up and down arrow keys so that the researcher could have some control over the map (if needed) as they maintained possession of the keyboard.

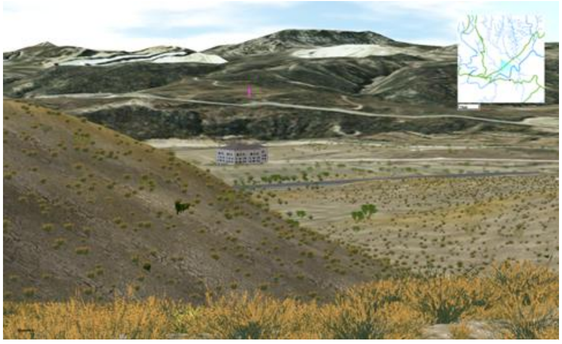
Researcher interaction was centred on the keyboard input device and was primarily used for system configuration and control of the experimental features (trial location and map type). Numbered keys were used to represent the different trial locations (1-3 for the first estimation and 4 for the last) with the 0 key being used to return to the default motion model if required. Display of the maps was controlled via character keys with 'n' being used to turn off the map, 'm' to show the static map, and 's' to show the dynamically scalable map. The 'd' key was also used to turn on or off the scale bar.

5.2.3 Experimental Design

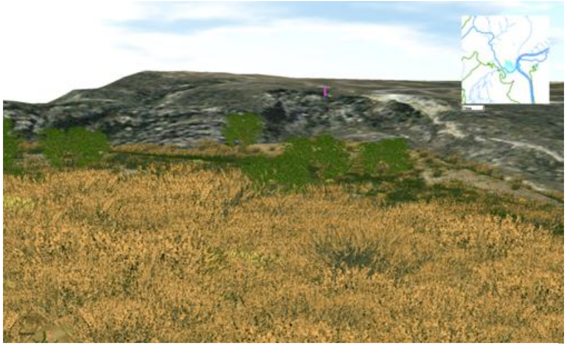
In this study, two aspects of distance estimation within VEs were to be investigated: whether providing an overview map effects accuracy of estimated observational distances; and whether having overview maps affects transfer of knowledge for estimating distances when they are removed.

In both investigations the user was asked to estimate the distance between their location in the environment and an abstract marker within the scene. For the first investigation, each user was placed in one of three pre-defined locations and presented with either no map, a static map, or a dynamically scalable map. In all three locations the horizontal distance between the observer's location and the marker was the same (approximately 1100m). Once an estimation was provided, the map (if present) was removed and the user's location within the environment moved to a second area and again asked to estimate the observational distance between themselves and another marker. This second location was the same for all test subjects, and the distance different from the previous task (although it was the same for all test subjects at approximately 1600m). The four locations can be seen in Figure 16.

For the experiment, three varying locations were used for the first distance estimation. The implementation of multiple locations was used as in each scene different features that could be used for estimating distances were visible between the map and environment. For example location 1 has more features present in the environment than appear on the map (see Figure 16) and so this could have an impact on the relative effectiveness of the map. Additionally, different features in general may have effects on estimations owing to personal preferences. For example, it may be the case that main roads and buildings could be more useful features to people who are accustomed to urban settings, and rivers and hills more useful for those with more rural environment experience. Additionally, someone who knows about geomorphology may be able to derive distance information from the general terrain easier than someone who is from an urban geography setting.

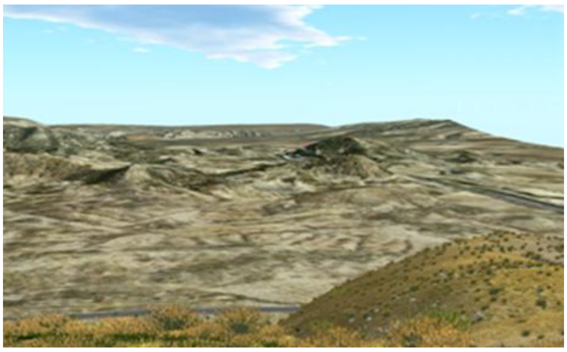


Location 1



Location 2

Location 3



Location 4

Figure 16: Trial locations

The locations of the markers were initially determined by first randomly selecting the viewing locations and then generating a buffer of 1100m around each of the points using GIS software. The marker location was then placed on a point located on the buffer perimeter and the scene viewed from the viewing location. In some instances the marker location was unsuitable owing to it being obscured by features such as trees and buildings or by the terrain itself. In those instances a new marker location was selected and then the suitability re-assessed. The actual as-the-crow-flies distance to the markers were not exactly 1100m however owing to the differences in elevation between the viewing and marker locations, as well as slight inaccuracies in the positioning of the markers on the buffer edge. Also placement within x and y directions was to the nearest whole metre. The reason for randomly selecting locations was to ensure that results obtained reflected the general environment rather than a particular subset. As it may be the case that particular features within a scene make the task easier or harder, the decision was made that the features present within the view should not be a sampling criteria as a sufficient knowledge regarding the individual impact of each item on the distance estimation was not known. Therefore a selection method which would cover all different aspects was not possible and so the random selection method was implemented.

Before all trials participants completed an informed consent form and were given the opportunity to ask questions. They were also given the opportunity to decline the use of audio recording equipment and not have their quotes used in any written work based on the trials. Ethical clearance was obtained from the University of Leicester Geography Department Ethics Committee before the trial portion of this investigation.

In addition, a short pilot study was conducted before the main trial phase as a means of determining the feasibility of several aspects of the trials. Initially a numerical scale bar was used (showing values such as 1:15,000) as opposed to the graphical format. During the pilot study, it was found that this delivery method was too difficult for participants to interpret and so the graphical method was developed. It was also decided from the pilot study not to implement any time constraints on the task, as well as to allow the user the option of having the researcher control movements within the environment as some people found the controls particularly difficult.

5.2.3.1 Distance Estimation between Map Types

As mentioned, the first distance estimation task was performed to investigate whether presenting the user with an overview map affects accuracy of distance estimations and if so, does the ability to zoom in have any effect. The different map types used were no map, static map, and dynamically scalable map. Within the task, users were asked to provide verbal estimations as accurately as possible (at least to one decimal point of a kilometre), although conversion between metres and feet could be undertaken if preferred by the test subject. No time limit was imposed on the task.

Actual distances provided were to be in the form of absolute egocentric distances, meaning that they were an estimation in the form of a known numerical standard such as metres or feet (Thompson et al. 2004) and are distances between the observer and object. Thompson et al. (2004) also describe two other forms of estimation: relative and ordinal, where relative is in terms of another distance (i.e. x is twice as far as y) and ordinal is again the distance in terms of another object, but without magnitude (i.e. x is farther than y).

For this experiment, three locations were used where a marker was placed at a horizontal distance of 1100m from the viewing location. Owing to the differences in elevation between the viewer location and the marker each distance was actually different as mentioned earlier (1101.5m, 1098.5m and 1101.1m).

5.2.3.2 Transfer of Knowledge Task

After an estimation had been given for the first distance, the subject was ‘teleported’ within the environment to a new location and asked to estimate again the distance to an abstract marker. In this section of the trial however the assessment conducted was on the effects of the maps on transference of knowledge, and therefore the map would not be displayed. Doing such a task would indicate whether having a map in a previous task had any effect on subsequent estimations when it was taken away, and also if the type of map had an effect. As the location and distance for this estimation was the same throughout all trials, any significant change between groups would be directly related to the map presented in the previous task.

5.2.3.3 Qualitative Data Collection

Qualitative data were collected from users via post-trial interviews and in-trial feedback. For the in-trial feedback users were asked to verbalise their thought processes when constructing their distance estimations. Post trial interviews were conducted in a semi-structured manner with a set series of questions which were then investigated further depending on the interviewees responses. Questions asked during the interview process were:

- What in the environment did you use to estimate the distances?
- Did the map help?
- Did the ability to scale the map help?
- Did you use experience gained in the first trial to help in estimations for the second one?
- How accurate do you generally think you were?

The purpose of these questions were to gain more insight into whether a specific aspect of the map was of benefit more than another and whether the user felt that the map or scaling ability helped in any instance. This was due to the possible outcome that either the user felt that the map did not help but quantitatively it did, or vice-versa. Additionally by asking about features used in the environment, locations that could be more difficult in terms of lack of features could be identified if needed and then identify if the map was of more benefit in locations was more useful in such areas or not.

5.2.3.4 Quantitative Data Collection

Quantitative data relating to distance estimations was obtained through verbal distance estimations where users were asked to estimate the distance to the marker to the nearest 100 metres. Distances could be provided in imperial measurements (feet, miles etc.) if the user preferred and then these were converted to a metric counterpart for analysis.

5.3 Results

27 Participants took part in the trials comprising of 10 males and 16 females (1 gender not recorded) with an average age of 28 (based on age bands using the centre of each band to determine average, with n=17 owing to error in data collection). Each participant was also given the opportunity to enter a prize draw to win a £10 book voucher as an incentive. Participants were predominantly students or staff within the geography department of the University of Leicester.

5.3.1 Quantitative Analysis

After the collection of numerical distance estimations from trials, conversions were undertaken to gain the absolute relative error values for each estimation. The implementation of absolute relative error was to give the unsigned error value to identify by how much the estimations were different from the actual distance between the map types. This made for better analysis using the Friedman Ranked Sum test which will be discussed shortly. Additionally, as the distances are slightly different between the locations a relative value was used which signifies the proportion of the actual distance that the estimated distance was out by. For example, if a estimation of 1200 metres was given for a distance of 1100 then the absolute relative error would be 0.0909 indicating that the estimation was 0.0909 times larger (100 metres) than the actual distance. If for a distance of 2100 metres a value of 2200 was estimated, then the absolute relative error would be 0.05 indicating that the estimation was 0.0476 times larger (100 metres). On the other hand if the user underestimated a distance of 1500 metres with an estimation of 1100 metres then the resultant absolute relative error would be 0.2667 (400 metres). Again, as we are interested in how much error occurred we do not need to know that it was an underestimation and so we use the unsigned relative error value. Additional analysis was conducted by assessing the normality of the signed data, although this was still performed on the relative values. Table 7 contains the raw data collected from the trials. From general observation the levels of over and under estimation appear to be balanced (actual distances are L1=1101m, L2=1098m, L3=1101m and T=1853m) with 13 over estimates and 14 under estimates in both tasks. Similar equality was found within the map type groups (Table 8). Quantitative analysis was performed between the test groups (determined by the location) and the test treatment (the map type) to identify any relationships between task performance and the use of an overview map, as will now be discussed.

Trial	Map Type	Location			Transference
		1	2	3	
1	N	1000	-	-	4000
2	M	-	2500	-	4300
3	S	-	-	1100	2100
4	N	-	-	1500	2000
5	M	3500	-	-	4000
6	S	-	1200	-	2200
7	N	-	805	-	1207
8	M	-	-	1200	1500
9	S	1100	-	-	800
10	N	2000	-	-	6000
11	M	-	1500	-	1300
12	S	-	-	500	1500
13	N	-	-	1200	2000
14	M	1000	-	-	500
15	S	-	400	-	2000
16	N	-	200	-	800
17	M	-	-	300	2200
18	S	5000	-	-	7500
19	N	1200	-	-	1600
20	M	-	2400	-	1600
21	S	-	-	1250	500
22	N	-	-	500	1000
23	M	800	-	-	700

24	S	-	500	-	1000
25	N	-	800	-	2400
26	M	-	-	700	1200
27	S	1500	-	-	2500

Table 7: Distance estimation raw data

	Task 1			Task 2		
	No map	Map	Scalable map	No map	Map	Scalable map
	-	+	-	+	+	+
	+	+	+	+	+	+
	-	+	-	-	+	-
	+	+	-	+	+	-
	+	-	-	+	-	-
	-	+	+	-	-	+
	+	+	+	-	+	-
	-	-	-	-	-	-
	-	-	+	+	-	+
Over estimate	4	6	4	5	5	4
Under estimate	5	3	5	4	4	5

Table 8: Distance estimation trial over and under estimates (+ signifies over and - signifies under)

5.3.1.1 Distance Estimations with an Overview Map

The data collected were analysed using a modified Friedman Ranked Sum statistical test developed by Brunsdon (2009), which henceforth will be known as the Replicated Friedman Test. The original Friedman Ranked Sum test works by determining whether a treatment has an effect on groups of data. It is a nonparametric test and suitable for small subject sample sizes (Darken 1996). In the case of this experiment the treatments are the map types implemented and the groups are the different locations. However, the standard Friedman test does not take into account multiple observations for the same treatment from the same group and so the test needed to be modified to allow for the replication of observations for each treatment. The method implemented was developed by Professor Chris Brunsdon using the R statistical package using the code which can be found in Appendix 1 (Brunsdon 2009). The Replicated Friedman Test statistic is given by:

$$\frac{p-1}{p\sigma^2}\sum_{j=1}^{p}(\bar{r}j-\bar{r})^2$$

5.10

where $\bar{r}j$ is the average rank over all replicates and all locations of map type j, $\bar{r}$ is the overall average rank, p is the number of locations and σ^2 is the variance of the ranks (Brunsdon 2012). The modified statistic calculates the average rank for each treatment (map type) across each group (location). Unlike the original Friedman test, the replicated test does not assume that there is only one observation for a treatment within each group. However, it is then possible that the theoretical distribution of the statistic given in Friedman's initial paper (Friedman 1937) may not hold, so evaluation of the statistic against a null hypothesis is achieved via Monte-Carlo simulation.

The Monte-Carlo simulation of randomly assigned locations and map types was performed for 10,000 repetitions. During each repetition the table of ranks is randomly calculated and the test statistic calculated. The final test-statistic is provided by identifying the percentage of times that the simulated test result from the Monte-Carlo simulation is greater than the one generated from the actual data. The resultant percentage value indicates how significant the results from the data collected are by showing the probability of the calculated test statistic being greater than that of one derived from the random generations in the Monte-Carlo simulation. If the p-value was very low ($p<0.05$) this would indicate some significance in the data.

Loc.	**Map Type**								
	No map			**Static Map**			**Scalable Map**		
1	0.092	0.816	0.090	2.178	0.092	0.274	0.001	3.539	0.362
2	0.267	0.818	0.272	1.276	0.366	1.185	0.093	0.636	0.545
3	0.362	0.090	0.546	0.090	0.728	0.364	0.001	0.546	0.135

Table 9: Absolute relative errors for distance estimation task

In the case of the data collected from this experiment, an end p-value was calculated at 0.655 indicating that there were no significant differences between the treatments of different map types (Figure 17). In addition, it was found that of the 27 estimations 14 were underestimates and 13 were overestimates. This indicates that there is no bias towards a general underestimate or overestimate in the distances provided.

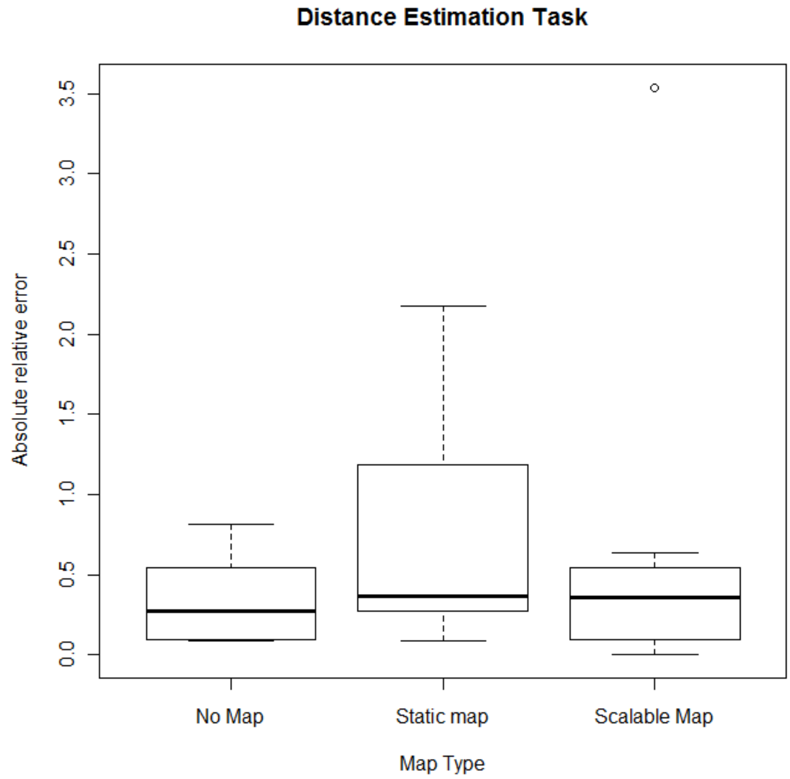


Figure 17: Boxplot of distance estimation absolute relative errors

5.3.1.2 Transfer of Knowledge

To determine whether the presentation of a map had any influence on later estimations where it was removed the data collected from the transfer of knowledge task were analysed using the same method as the main experimental analysis. The absolute relative errors collected can be seen in Table 10.

Using the same replicated Friedman test discussed earlier a p-value of 0.282 was obtained indicating that there was no significant differences present between any of the map presentation types (Figure 18). Again, it was found that there were 13 underestimates and 14 overestimates in the distances provided indicating no bias towards a particular direction of error.

Loc	Map Type								
	No map			Static Map			Scalable Map		
1	1.159	2.238	0.137	1.159	0.730	0.622	0.568	3.047	0.349
2	0.349	0.568	0.295	1.231	0.299	0.137	0.187	0.079	0.460
3	0.079	0.079	0.460	0.791	0.187	0.352	0.133	0.191	0.730

Table 10: Absolute relative errors for transfer of knowledge task

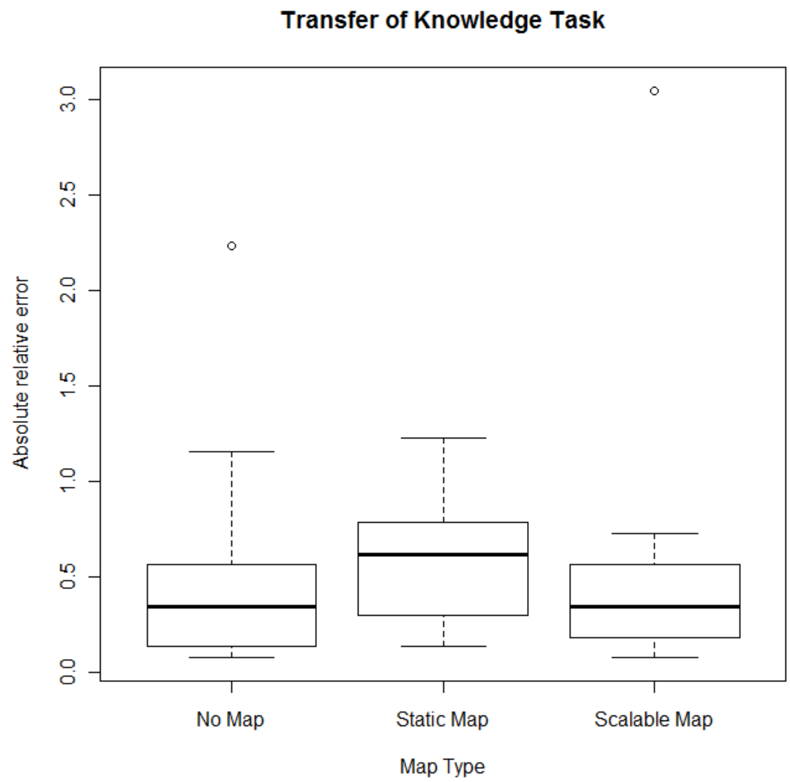


Figure 18: Boxplot of transfer of knowledge absolute relative errors

5.3.1.3 Testing by Location

As well as performing tests based on the map type being used as the treatment, analysis was also undertaken where the locations were seen as treatments to investigate if there were any significant differences between the estimations given in different locations. When analysing the first estimation we find no significance from the statistical analysis (p=0.678) and visual analysis of the data via boxplots (Figure 19) shows that the distributions look fairly similar although two outliers are identified in the location 1 data.

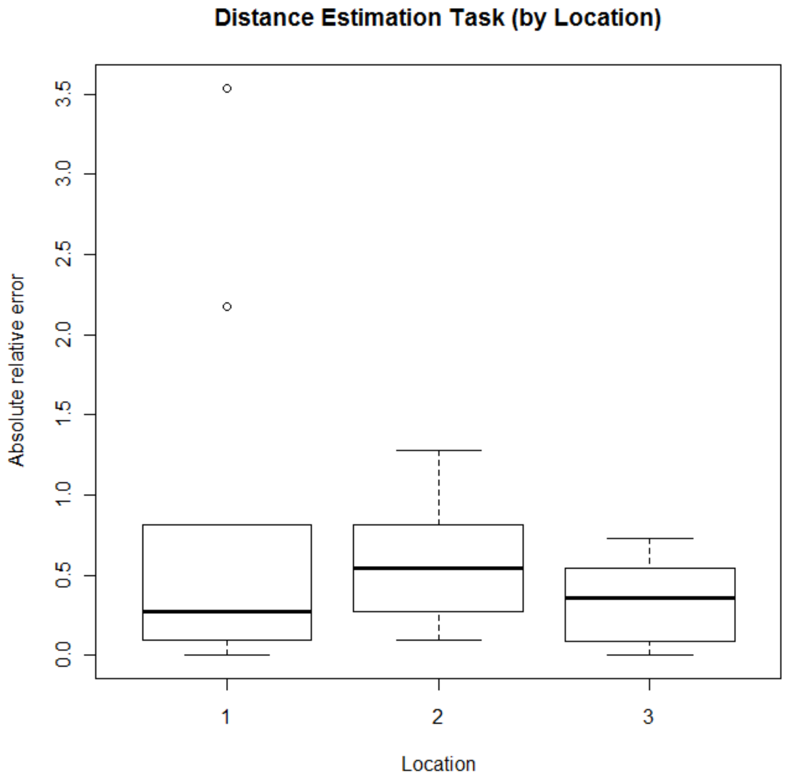


Figure 19: Boxplot of distance estimation errors (by location)

When we do the same analysis for the transfer of knowledge tasks with the location used as treatment we again find no significant differences but the p-value is much higher at 0.924. Although this does not identify any significance, when looking at the boxplot of the data (Figure 20) we see that the absolute relative error generated from where the users were located in position 1 for the first estimation appears to be higher in the transfer of knowledge tasks. Again it should be emphasised that there is no significant differences from statistical analysis in the current investigation, although further investigation may be warranted owing to the apparent influence of the location. This is however outside of the scope of this study.

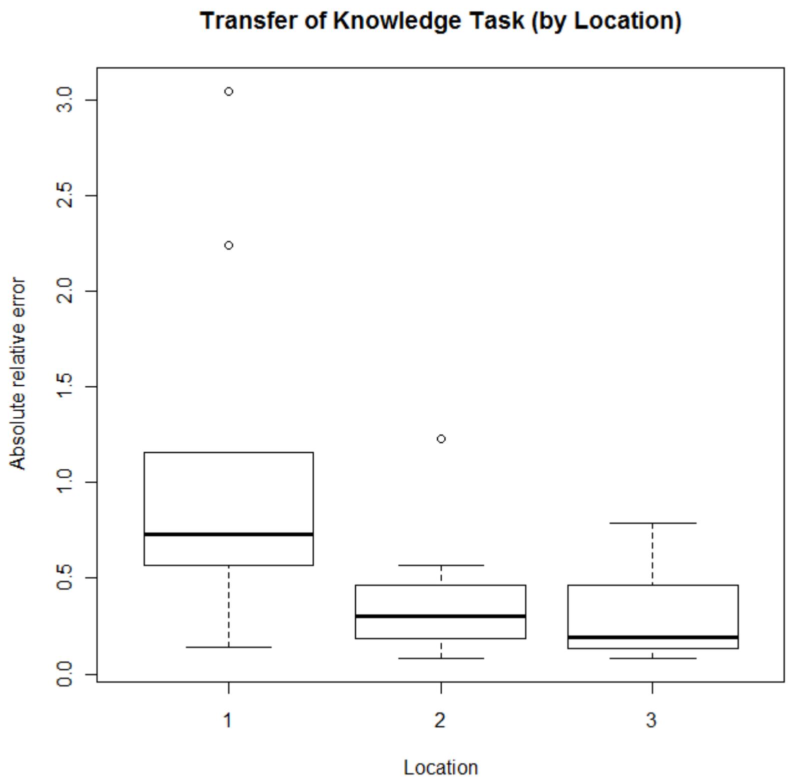


Figure 20: Boxplot of transfer of knowledge errors (by location)

5.3.1.4 Performance between Tasks

As a final piece of quantitative analysis performance between the first and second estimations was assessed to identify if there was a change in performance between the two tasks. Using paired t-tests to compare the estimations between the first and second tasks it was identified that there were no significant differences between the two estimations (p=0.152, 0.129, and 0.719 for the no map, map, and scalable map groups respectively). This indicates that the test subjects did not find one task more difficult than the other.

5.3.2 Qualitative Analysis

Data was extracted from the interviews and in-trial feedback by transcribing the audio recordings and identifying key aspects. Specific questions were used to highlight characteristics of the estimation process with in-trial feedback being used to assist in the identification of features used and whether the distance given in the first estimation was being used as a guide for the second.

5.3.2.1 Was the Map Used?

All test subjects who had a map displayed were asked if they made use of the map presented to them. Of the 18 trials where this was the case, interviews identified that 13 people made some use of the map to determine the distance, with the other five stating that they did not use it at all.

5.3.2.2 Was the Scalable Function of the Map Useful?

The test subjects who had the ability to scale the map through user interaction were asked as to whether this functionality was useful in the estimations. Of the nine participants where this was the case, eight said that it was useful and only one stated that it was not. It should be noted however that one of the test subjects who said that it was useful did not actually use the map to derive their estimation. Instead they identified that it would have been useful if they were to have based their estimation on it.

5.3.2.3 Which Estimation Did Users Feel was Easier/More Accurate?

As discussed, the trials consisted of two estimations: one where a map was displayed to the user based on the trial number (either a scalable map, a static map, or no map at all); and one where it was always missing. Some responses in interviews identified that users felt that one estimation was either easier or more accurate than the other. Of the 17 users who made such remarks six felt that the second was easier or more accurate and 11 felt that the first was easier or more accurate. Within the six users who felt that the second was easier or more accurate, five undertook trials where no map was present and within the 11 who felt that the first estimation was easier or more accurate only one was from a trial where no map was present. When the second one was stated to be easier, responses included

> *"Second one was easier because I had a benchmark from the first one for me to go further"*

> *"a bit of a learning curve"*

and when the first one was stated to be easier responses given were

> *"... fairly good on the first one with the scale"*

> *"because there was more information to go off from the picture [environment]"*

> *"I had the map to help me out".*

5.3.2.4 Was the Distance From the First Estimation Used as a Guide in the Second?

By analysing feedback during the trial and responses to interview questions, information was extracted with regard to whether the distance used in the first estimation was somehow used to help in the determination of the distance in the second. From the 27 trials, 12 users stated in interviews that the first estimation had an influence on the second, nine said that it did not and five implied that it did from the in-trial feedback. When users implied that the first had an influence responses included:

> *"Right, that's further away, that is nearer ¾ of a mile."*

> *"That to me looks a little bit closer."*

5.3.2.5 What Features Were Used to Estimate the Distance?

From feedback during the trials and questioning during the interviews features that were used to aid in the estimation of distances were identified. These features can be seen in Table 11.

From this information it is clear that the predominant features used was the topography of the environment, natural objects such as bushes and trees, and manmade features such as buildings and roads. When a map was present this was also identified as being of use in the majority of cases.

Feature type	Example	Count
Buildings/Roads	*“so I would say the house was about half a mile away”* *“the road looked really close”* *“the roads are usually a certain width”*	10
Layering	*“Yes it does [stereo helped] because it’s easier to identify apparent layering.”* *“noting that there was stages in it so you’d have a foreground, a middle ground and a further in the background”*	2
Map	*“I’m probably judging it more by the map”* *“For the first one I looked at the map and tried to estimate how far away it was by looking at the map”*	13
Natural Objects	*“Ok, that shrub is that big “* *“This rock is like 50cm or something”*	14

Past Experiences	*"I'm sort of thinking back to my California trip if you looked all the way across to the other mountain it might not of been that far away but on a map it's 50 miles away."* *"I actually found myself thinking that if it was a golf hole over there how far would I think it was?"*	9
Texture Compression	*"In the second one I can see three effective resolution levels"* *"Well maybe the, what's the word, texture [...] but when it's far away it's not made of grass, as you can see here in the foreground."*	2
Terrain/Topography	*"I looked the way they were disappearing away off in the first ravine and you could see the ravine come back up again and I was just estimating based on that."* *"But also looking at how big the terrain looked"*	18

Table 11: Estimation features

5.4 Findings

Using the analysis of both qualitative and quantitative data a general overview of how the presentation of overview maps affects the estimations given for distances within a VE has been conducted. In the following subsections the research questions identified are answered.

5.4.1 Does Providing an Overview Map Inside a Virtual Environment Aid in User Estimation of Observational Distances?

Using the estimations collected from the user trials, statistical analyses were performed which identified that the presentation of overview maps in either the static or scalable format had no significant impact on the estimations provided. However, the results obtained from the interviews indicate that when the maps were presented to the users they felt that the task was easier, or that they gave more accurate estimations than in subsequent estimations when it was removed. When no map was present in the first estimation, users felt subsequent estimations were more accurate or easier to make. This seems to imply that although users were statistically no more accurate when the map was presented they felt that they were. Previous studies have discussed that it can be difficult to derive a mental representation of a 3D landscape from a 2D map (i.e. Dolman et al. (2001) and Appleton & Lovett (2005)). In the study by Appleton & Lovett (2005), the 3D transformation was from turning a 2D planning proposal map into a 3D mental image of what the site would look like. In the study presented here the transformation was to overlay the information from the map to the already visible 3D representation, and therefore this may show evidence that the problem lies in general interpretation of the information rather than the conversion between dimensions.

To answer the research question posed, the provision of an overview map does not aid in the accuracy of observational distance estimations, but it does appear to make the task seem more approachable for users. In addition, no evidence was found of difference between whether the users had the ability to scale the map or not.

5.4.2 Does Providing an Overview Map in One Instance Improve User Estimations of Distance in a Second Instance When the Map Has Been Removed?

From the statistical analysis of the estimations given in the second estimation it is evident that having an overview map presented to the user in a prior estimation has no effect on the accuracy of subsequent estimations.

By looking at qualitative results from the interviews and in-trial feedback we see that in many instances (n=17, 63%) users identified that information about the distance in the first estimation was used as a gauge for the second. In some instances this information was used as an explicit measurement with responses including

> *"it's obviously a good 500 to 1000 metres beyond where the last one was."*

In other instances it was more of a relative measurement with feedback including

> *"I really don't have a clue, it looks further than the last one."*

From this feedback it is clear that distance estimations appear to be influenced by previous estimations. Therefore if the first estimation can be made to be more accurate then subsequent ones should also be more accurate.

In answer to the research question, providing an overview map in one instance does not improve distance estimations in a second instance when that map has been removed, but in several instances it was implied that that information was used relating to the first distance estimated.

5.4.3 Other Findings

As well as analysis of results to answer the primary research questions, additional information and analysis was obtained from the data collected.

When asked what features were used to help in their estimations the predominant response was the terrain and topography of the environment. Other aspects were also identified including the vegetation, buildings, roads and the map. Additionally, items such as layering and texture compression were also suggested by users as well as making use of previous experiences in real world instances.

Another interesting feature (although not statistically significant) within the transfer of knowledge task is when the users estimated their first distance at location 1. From the boxplot produced (Figure 20) it appears that when the first estimation was at location 1 the subsequent estimation in the transfer of knowledge location was more erroneous. When we compare the scenes (Figure 16) a possible explanation of why the difference would occur in estimations may be found in the location of the building in scene 1. When attempting to determine the distance in this scene it appeared that the building was often used as a means of determining a ‘half distance’ value which was then used to identify the distance to the marker. Additionally there are more features within the scene (roads) which coincide with features shown on the map. Owing to the lack of such features in the transfer of knowledge location the half distance skill and comparison between features where a distance could be calculated on the map could not be implemented, and so if these were relied on more in the first estimation then the skills obtained would not be transferred. In the other locations where aspects such as the terrain were used more, the similarities between the initial and transfer locations allow for the transfer of the skills gained in using these.

5.5 Summary and Conclusions

Although the results of this experiment have found no statistical evidence of an effect on distance estimations by the presentation of an overview map, they have highlighted an important consideration. It has been found that although there is no effect on the estimations given, users tended to feel that their estimations were either more accurate or easier to do when the map was shown. This means that the users were more confident in estimations that were just as erroneous as when no additional aid was presented. In this instance this does not cause too much of a problem but in real world instances this effect could be far more serious. For example if someone is navigating to a location within a rural environment and they know they only have an hour of daylight left then they have the choice of either going somewhere closer for shelter or go to where they need to be. If they are unsure about how far away the end point is then they would be more likely to seek shelter in the more immediate area as they would not be sure as to how long it would take to get to the end point. If however they used a similar method to estimate the distance as presented here (i.e. an overview map on a mobile device) then they would be more confident in their estimation and so more likely to set out to the end point if they deem it is within walking distance in the time left. As shown in this study though, they may still be as inaccurate in their estimation as if they had not used the map and so run the risk of not reaching their destination before night fall.

However, the non-significant results from this study could be due to a number of reasons. Firstly, they could be simply due to the relatively small sample size. If a larger sample was used then it is likely more well defined distributions would be obtained which may identify distinct groups between the map implementations. Also, there may be confounding variables present in the study which affect how the map can be used for estimating distances. Such variables could be the design and interaction method for the map itself, or effects introduced from the environment. It would be difficult however to identify and address all of these variables without a thorough investigation into a number of aspects relating to the task, map and environment. As such, an investigation into these was not possible in the timeframe provided.

The other primary aspect investigated in this study was whether having an overview map in one instance had an influence on a later estimation when it was removed. Again statistical analysis has shown no evidence of such an influence, although the qualitative data collected does indicate that when a map was present in the first instance users found that estimation easier than when the map was removed, and that when no map was present at all users found the second estimation easier.

As a conclusion, it is clear from this study that although the presentation of the map to the user has not had any effect on the estimations provided it does make users feel that the estimations are easier to determine and that they are more accurate. As discussed, in real world applications this could cause problems and so care is needed when using overview maps in such tasks. With regards to transfer of knowledge again no evidence was found of effects caused by the presentation of a map in the first estimation. However, visual analysis of the boxplots produced from the transfer of knowledge data suggests that there may be a difference in estimations between the locations used in the first estimation task, although no significant results were found. In both sets of estimations, there was no apparent bias towards either underestimation or overestimation of distances.

5.6 Further Study

One of the questions arising from this study is how the environment affects the transfer of knowledge from one estimation to another, in the case of this study between an estimation where an overview map is present and one where it is not. Although no statistical evidence of effect was found, indications from the data may suggest that in further studies which directly investigate the transfer between locations a difference could be found. Therefore it would be of benefit to investigate this aspect of the experiment in more detail, possibly investigating specific feature presence, land cover types and general topography of the environment. In this study several features that were used to help estimate distances were identified, and these features could be used as a basis for identifying differences between scenes.

Another study which would be of benefit would be to investigate in depth why the maps had no influence on the accuracy of the estimations. For instance would a different map design or interaction method alter the result found here, or possibly emphasis on the scale bar could make it easier to use. As it was believed that the presentation of the scale bar and map information would logically improve estimations, would a different implementation show this or would it come up with the same results presented here?

6 Part II: The Effects of Info-Marks on Performance in a Route Tracing Exercise

Does the presentation of areal geo-located textual information affect navigational performance in a route tracing task?

- ***Does the method of presenting info-marks make a difference to their effectiveness?***

Which features in a virtual rural environment could take the places of the five feature types of the Urban Image Theory?

6.1 Introduction

The second part of the study was to look at another aspect of spatial cognition in VEs – the process of orientation and navigation. As discussed earlier, these two processes are closely related. Orientation is the determination of where you are in an environment and navigation is the process of getting from one place to another. Obviously to navigate successfully you must be able to determine where you are at key points along your journey.

One method of determining your location is to use distinctive features in your vicinity, which are known as landmarks. By determining their position in relation to yourself you can pinpoint your location in the environment. Earlier discussions showed that these landmarks need to be unique to be useful and that obviously the more there are the more accurate the orientation process will be. However in some environments such as large scale rural areas, the dispersal of features that can become landmarks within the region is increased. This not only reduces the number of landmark features but also their inter-visibility. Therefore it makes sense that in such locations the ability to orientate oneself will be diminished and so tasks requiring navigation in the environment will suffer. This makes it clear that a method of providing landmarks to the navigator is required that not only increases the number of landmark features present in terms of density, but also allows their usage even if they are not necessarily visible.

The approach implemented in this investigation was to provide landmarks on an areal basis, meaning that rather than the landmark always being a particular visible feature, the presentation of the landmark information occurs whenever the person is within a specified proximity of it. For example, an aesthetically pleasing building in an urban environment may be a very useful landmark for navigational purposes as it would stand out from the surrounding structures. However, if the surrounding structures are very tall then it's usefulness would be reduced as it could very well be the case that the landmark building could only be seen when you are in fact standing in front of it. It is still useful in terms of you know where you are when stood in front of it, but as a more global landmark it is not useful as it can only be seen from this one view point. If you could see through all the buildings around it however, or even if it created an 'aura' around it that altered the perception of surrounding features then it could be used as a global landmark. Therefore, one approach of making it more useful would be to let the navigator know that they are within range of it even if they cannot see the landmark feature itself.

This study made use 'info-marks' in the form of textual information about specific features or locations that were presented to the user via either an onscreen display or a hand held PDA device. This presentation of information occurred when the user entered a region within one kilometre of the location/feature in question within the VE. The PDA display is achieved by the system transmitting a Bluetooth signal that imitates that transmitted by a GPS receiver. This signal indicates the GPS location as being the user's location within the VE as opposed to their geographical location in the real world, as implemented by Priestnall & Polmear (2006).

Another aspect investigated here was with regards to the UIT. As discussed earlier, Vinson (1999) stated that features from all of the UIT classifications should be present in the environment to make it more navigable. However in non-urban locations the features falling into these classifications are less well defined, and so in this study the users were asked if they observed any features that could fit into the classifications. By identifying such features in a rural context, later environments can either make use of the UIT by highlighting such features or by including more of them if possible.

This chapter describes the methods of presenting the info-mark data to the users detailing their creation, determination of which info-mark to display, and the methods of delivery. Analysis is undertaken using data obtained from track logs as well as qualitative data from interviews and in-trial feedback. Finally the results obtained are used to identify whether the presentation and delivery method of the info-marks had any effect on the user's performance with regard to following a route shown on a paper map, and if the delivery method of such information altered their effectiveness. The chapter also describes the use of qualitative analysis to determine the placement of features within the environment into the different Urban Image classifications. Also presented are possible improvements to the environment along with aspects that warrant further study.

6.2 Methodology

The same Blueberry model of the Sorbas region developed for the previous experiment was used again for this investigation. However, some modifications were made to the original model by incorporating more recently acquired DEM and landcover classification data, although the same Blueberry landcover layers (visualisations) were applied to this new dataset. This meant that the actual look of the environment did not change extensively, but the land cover areas were more accurate in their resolution and location as the data obtained provided more information than the original dataset. A second completely artificial environment was also created as a training ground, which allowed users to become accustomed to the controls before beginning the trial.

Additionally, a different motion model was implemented in this system, making use of VegaPrime's gaming motion model as opposed to the default UFO motion model. In VegaPrime the motion model is the method the user moves and interacts with the environment, and the UFO motion model is the name given within VegaPrime for the simple motion model that allows flying through an environment. There are several different models which are each aimed at different in-environment movement techniques such as aircraft, walking and driving. In each of the models the user input from keyboard and mouse is interpreted and then portrayed in the environment as movement. The default UFO system uses mouse actions to move forward and backwards, look around and alter the speed. Although this is the simplest motion model a major issue is that the view continues to move even if the mouse is stationary (unless the mouse pointer is returned back to the centre of the window). This has shown to be problematic for non-experienced users. In the gaming motion model, a user holds down the left mouse button and moves the mouse to look around, and uses the arrow keys to move. This made movement easier for the test subjects as once the mouse was stationary the rotational movements also became stationary as opposed to the UFO motion model where they continued to move. This was important as the users would be controlling the movement within the environment and many would not be experienced with moving around a VE.

Further functionality was applied to the system to display the info-marks onscreen and to allow the transmission of a Bluetooth signal imitating that which would be sent by a GPS receiver (the signal being used by a PDA running software to display the info-marks). Datasets were collated for the info-marks, the method for storage and determination of display dependant on the interface type (onscreen or via the PDA). The system was also updated to allow the logging of user movements, recording a track log throughout the trial process.

Analysis was performed by extracting data from the track logs and post-trial interviews. Brunsdon et al. (2007) state that when assessing real-life route finding, it is important that analysis is not performed on a single unitary measurement and should instead involve multi-dimensional analysis on a range of outcome measures. Therefore, track logs were analysed using a series of metrics involving time taken, distance travelled, average speed and final distance from end point.

6.2.1 Construction of the Info-Mark Dataset

The areal landmarks used for this experiment were not based on highly visible features – in fact, some of the landmarks used were not visible at all (i.e. underground cave systems). Instead, interesting features or important locations in the area were used and then feedback of their presence provided by text based 'factoids' presented to the user which are henceforth called info-marks. The reasoning behind the use of features that were not always visible was to ensure that there were info-marks present in a large proportion of the environment. In the case of this study, the point location of the feature (whether visible or not) is presented on a map so that a geographic location can be determined at all times. If the environment was to be learnt with the info-marks present then it is likely that the user would be able to use combinations of other unremarkable yet visible features in the environment as a means of anchoring the location of the feature that the info-mark is in relation to (i.e. a combination of background features).

6.2.1.1 Sources of Data

The data used for these 'info-marks' was collected from various sources such as geological books, websites and visible observations. Care was taken to attempt to disperse the info-mark dataset across the environment to provide information in as many locations as possible. The information displayed to the user for each info-mark can be seen in Appendix 2, with their locations displayed in Figure 21.

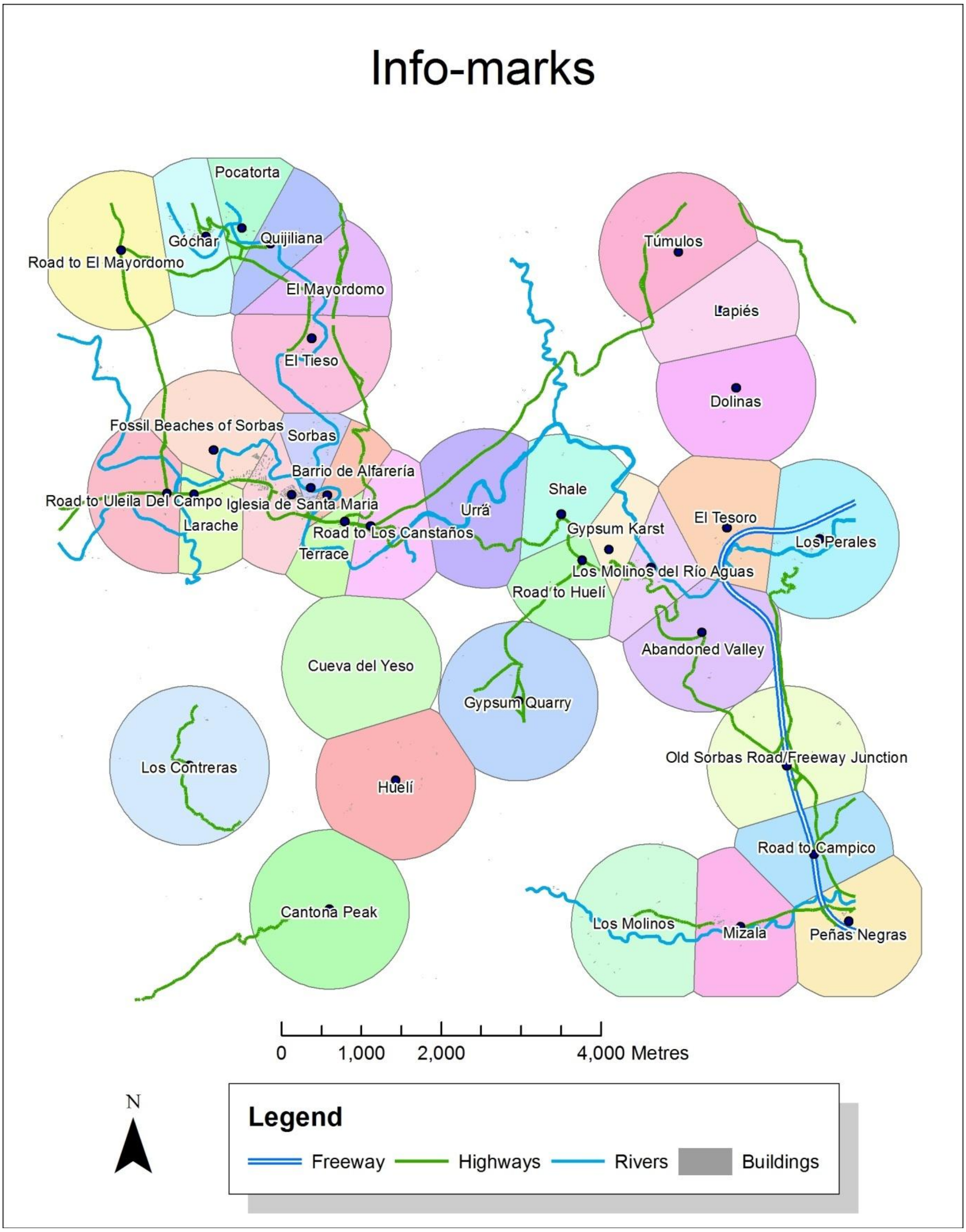


Figure 21: Info-mark locations

6.2.1.2 Creation of the Dataset

Once information had been collected with regards to the text being displayed to the user, Arcmap was used to create the vector data required to determine when each piece of information would be presented. A point shapefile was used to store the geographical location of each info-mark, with the coordinates being determined either by current point datasets (for villages), road vector intersections (road junctions), or transference of point/area information from reference documents.

As the display of the info-marks is based on proximity, a buffer of 1km was placed around each point feature. To ensure a suitable transition boundary between two info-marks, Thiessen polygons were also created for the point dataset, and then the buffers masked against these. The complete process used was as follows:

1. Create a point feature shapefile and add points for each feature that is to have an info-mark assigned to it.
2. Create a 1km buffer around each point, choosing not to dissolve the boundaries and storing the original point feature id against the polygons created.
3. Create Thiessen polygons for the point dataset, again ensuring that the point feature id is stored against the corresponding polygon.
4. Perform an intersection process between the buffer and Thiessen polygon datasets.
5. Select all features from the resulting intersect dataset where the point ID from the buffer polygon is equal to the point ID of the Thiessen polygon.
6. Dissolve the selected features based on one of the point Ids (from the Thiessen or buffer datasets).

6.2.2 Presenting the Info-Mark Data

As well as assessing whether the presentation of info-marks has an effect on the navigational performance, this study also investigates the effects between two delivery methods – virtually inside the VE display, and externally via a PDA device. Each of these presentation methods uses the same info-mark dataset but a different format for the areal aspect is required (Figure 22).

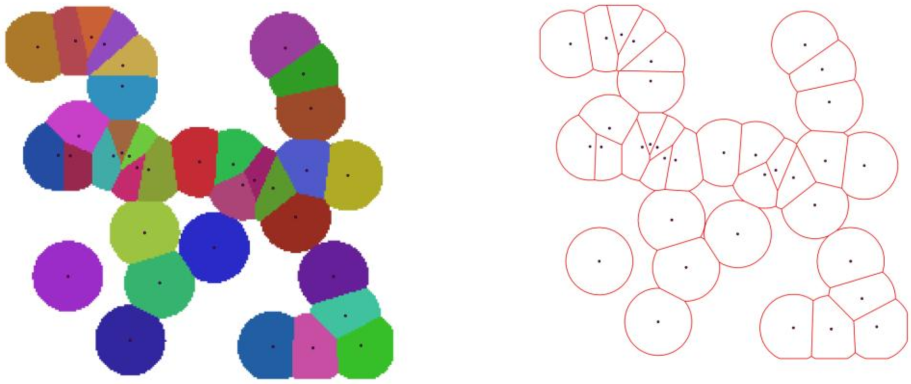


Figure 22: Info-mark dataset (left - raster, right - vector)

6.2.2.1 Presenting Info-Marks on a PDA

For the PDA method the vector polygons are used within a mobile based content delivery system called MScape.

MScape is a piece of software developed by Hewlett Packard which displays content to the user on a PDA dependent on readings from sensors attached to the PDA being used. The readings used to determine display include information from GPS, RFID tags and Infrared beacons, with conditional logic implemented via a scripting language. Information given to the user can be in the form of text, web pages, images, audio and video. In the case of this study the MScape software made use of the GPS signal, although the signal that the PDA receives indicates the location of the user in theVE rather than their real world geographic location (see section 6.2.3.3 below).

The MScape software used for presenting the info-marks to the user via the PDA device uses XML data structures to store information about the polygonal areas which dictate when to display the information. Using a conversion tool[7] the polygon data was directly exported into an MScape file ready for the inclusion of conditional logic to enable the display of the info-mark dataset. An example of the information presented to the user via the MScape system can be seen in Figure 23.

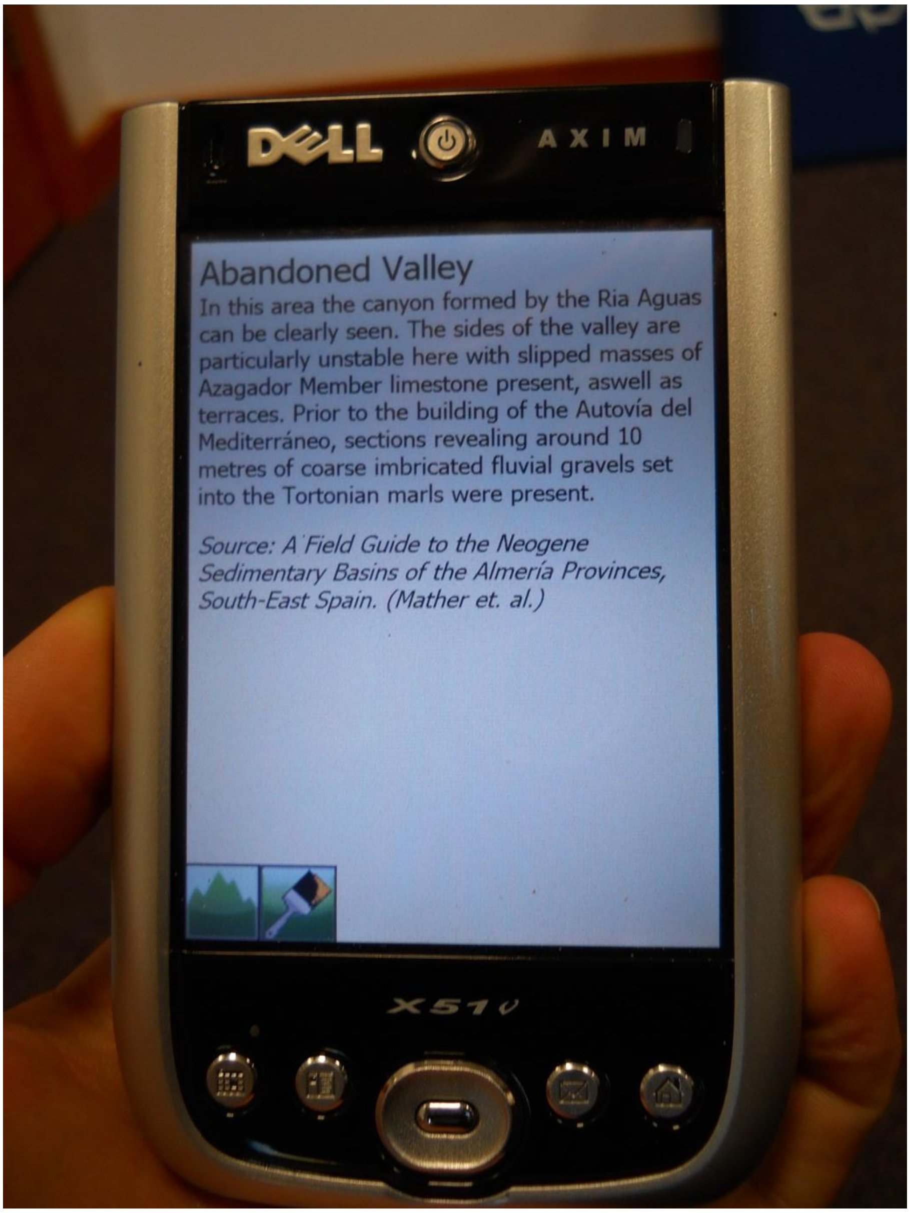


Figure 23: PDA info-mark delivery

7 Tool developed by Dr Jing Li, University of Leicester

6.2.2.2 Presenting Info-Marks in a VE

For the onscreen method, an ASCII file is read by the program to determine which info-mark (if any) is to be displayed at the current location. The reasoning behind this was that the values stored in the ASCII file could easily be read into the program and stored in a two dimensional array, and then the determination of information to display can be simply identified.

Conversion of the vector polygons to ASCII file was performed by first converting the polygon shapefile to a raster dataset using the ArcMap Spatial Analyst tools. The cell size was set to 50 metres to allow for a smaller file size (187Kb) and so faster access with less resource usage. In the case of this environment, the creation of a two dimensional array of size 219x210 was required, consisting of a total of 45990 cells with each cell holding an integer value requiring 4 bytes of memory. This then equates to a total of 183960 bytes (180Kb) needed to store the array in memory. As the visual aspect of the environment already puts heavy demands on the system it was felt that it would be better to have a smaller array size and thus reduce some of this overhead. Additionally it was decided to use integer variables as opposed to short integers to allow for better readability of code for future developments, as well as allowing for larger ID values if so required. Once converted into the raster file with cell values representing the info-mark ID, an inbuilt tool was used to convert the GRID dataset into an ASCII file.

The actual display of the info-marks on screen was performed via the texturing of a geometric primitive on the display. This primitive was a rectangle in shape whose texture was determined dynamically via information stored in the info-mark ASCII file created in ArcMap. The textures used were JPEG images created from the textual info-mark dataset used within MScape on the PDA devices, created simply by capturing a screen shot of each HTML file being displayed in a browser window of fixed size. These screen captures were then cropped in image editing software to remove the browser window frame and saved as JPEG files, with filenames of their info-mark ID.

Code was implemented into the system that read data from the ASCII file created from the info-mark polygon dataset in ArcMap (a sample of an ASCII file can be seen in Figure 24). To begin with, the header of the file was read to determine the dimensions of a two dimensional array used to store the ASCII file data and to identify the sizing of each cell. Once the command had been given to the system (via a key press) to display the information on screen, the current location of the user was used to determine which cell of the dataset corresponded to their location. This was done by using the lower x and lower y coordinates of the ASCII file and the cell size. For example, if the user was located at (578012, 4100312) then the following would be used:

$$Ax = \frac{(Cx - Lx)}{s}$$

$$Ay = nR - \frac{(Cy - Ly)}{s}$$

6.1

Where:

Ax, Ay = the x and y cell references

Cx, Cy = the x and y coordinates of the user in the environment (their Easting and Northing)

Lx, Ly = the lower x and y coordinates

s = the cell size

nR = the number of rows

This would then equate to an array reference of [205,62] (in two dimensional arrays, the row identifier comes first, and so in this case the y value). Owing to the nature of two dimensional arrays having their origins at the upper left corner and ASCII files having their origin at the lower left corner, the computation was needed to subtract the row reference value in the ASCII file from the total number of rows, this giving a reference value with origins at the upper left corner.

```
ncols         219
nrows         210
xllcorner     574874.21095848
yllcorner     4100047.4371738
cellsize      50
NODATA_value  -9999
-9999 -9999 -9999 1 1 1 1 1 -9999 -9999 -9999 -9999...
-9999 -9999 1 1 1 1 1 1 1 2 2 2 2 2 -9999 -9999...
-9999 1 1 1 1 1 1 1 2 2 2 2 2 2 2 -9999 -9999 -9999...
1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 -9999 -9999 3 3 3 3...
```

Figure 24: ASCII file structure

Once the array reference was determined, the corresponding value was identified. If this value was the no data value indicated in the header file, then the display placeholder had its visibility hidden. If however it was any other value, then this indicated that the corresponding info-mark needed to be displayed. In this instance, the JPEG texture file created from the HTML representations was loaded into the system and displayed on the primitive as a texture. Figure 25 shows a screenshot of an info-mark being displayed onscreen within the environment.

Figure 25: Scene from the environment showing an info-mark display

6.2.3 Environment Extension

As well as developing the info-mark dataset for display on an external device and developing code to allow for the display of info-marks onscreen, additional functionality was developed into the Sorbas Blueberry model to provide a compass, log user movements and to transmit a Bluetooth signal imitating a Bluetooth GPS receiver to the PDA.

6.2.3.1 Compass

To aid in the navigation process a compass display was implemented on screen (Figure 26). This compass comprised of a triangular needle pointing on a compass face (showing symbols for north, south, east and west) in the direction that the user is facing. The compass mechanism was implemented by displaying the face texture on a geometric primitive with the needle being drawn using an OpenGL triangle with the rotation matrix being updated by the yaw variable obtained from the system (the direction in which the user is facing).

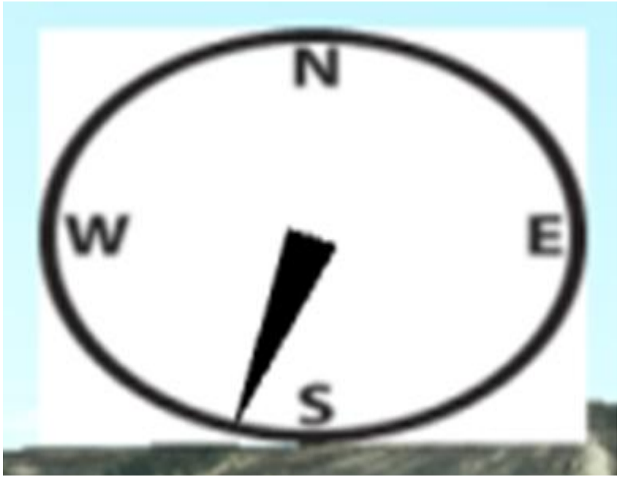


Figure 26: Compass display

6.2.3.2 Logging of Movements

To analyse the user performance during the trials, a method was needed to keep a record of the user's location and orientation throughout their attempt of tracing the route in the environment. As the environment itself and all movements are controlled by the program, it was possible to develop a simple data logging method to record the user's positional information throughout the trial. This method created a text file whose path and name were dependent on user input when the environment was first loaded and contained a structured list of the users x, y and z position, as well as the yaw, pitch and roll rotational values.

The value stored for the geographic location is in regard to the global origin located at (x=0,y=0,z=0) meaning that a stored value of (x=10,y=50,z=30) would indicate that he user was 10 units east, 50 units north and 30 units above the 'sea level' plane. The values stored for the rotations are in respect to the observer and not the global origin. It should be noted however that a rotational value of 0 indicates that the user would be directly aligned with the global axis if the user was at the global origin. For example, a rotational set of (y=0,p=0,r=0) would indicate that the user was looking due north parallel to the ground plane and in a non-tilted position. Values of (y=-45,p=15,r=20) would indicate that the user was looking 45° west of North, looking up from the ground by 15°, and tilted to the left by 20°.

An entry was made into a text file every frame refresh meaning that later analysis could select the level of detail with regards to the time duration between each sample point. It should be noted that the number of entries per second varied owing to the nature of the software used in that when a heavy processing load was placed on the computer then the number of frames per second being displayed dropped in turn meaning the number of entries into the track log per second also dropped.

6.2.3.3 Bluetooth Transmission

MScape can use the results from many different sensors to determine when to display specific pieces of data to the user, including positional information obtained via GPS. In the case of this experiment, it was required that the GPS data used by MScape was not indicating the PDA's real world location, but the position of the user in the VE, as performed by Priestnall & Polmear (2006). To accomplish this, functionality was added to the system that transmitted a Bluetooth signal in the form of a GPS sentence to be received by the PDA. MScape would then interpret this signal as being from an actual GPS receiver and as long as the sentence was in the correct format would use the information provided as an indication of its position.

The actual transmission of the location within the VE can be split into three main parts:

1. Determination of location in terms of latitude and longitude

2. Construction of the GPS NMEA sentence
3. Transmission of the sentence over a Bluetooth connection.

The internal coordinate system of the Blueberry model and VegaPrime environment is in Eastings and Northings using a Universal Transverse Mercator (UTM) projection. An NMEA GPS sentence however transmits positional information in latitude and longitude coordinates, and so an algorithmic function was created to perform the conversion using standard equations.

To ensure functionality between GPS receivers and devices, several 'sentences' (a structured sequence of characters) have been defined. The structure used for the transmission implemented here is the NMEA GPGGA sentence which transmits location data as well as information about the number of satellites and fix quality (general accuracy). The structure of the sentence is as follows:

```
$GPGGA,123519,4807.038,N,01131.000,E,1,08,0.9,545.4,M,46.9,M
,,*47

Where:
```

`GGA`	Global Positioning System Fix Data	
`123519`	Fix taken at 12:35:19 UTC	
`4807.038,N`	Latitude 48 deg 07.038' N	
`01131.000,E`	longitude 11 deg 31.000' E	
`1`	Fix quality:	0 = Invalid
		1 = GPS fix (SPS)
		2 = DGPS fix

	3 = PPS fix
	4 = Real Time Kinematic (RTK)
	5 = Float RTK
	6 = estimated (dead reckoning)
	7 = Manual input mode
	8 = Simulation mode
`08`	Number of satellites being tracked
`0.9`	Horizontal dilution of position
`545.4,M`	Altitude above mean sea level (Metres)
`46.9,M`	Height of geoid (mean sea level) above WGS84 ellipsoid
`(empty field)`	Time in seconds since last DGPS update
`(empty field)`	DGPS station ID number
`*47`	Checksum data (always begins with *)

Figure 27: GPS sentence structure

For the sentences produced by the system, the only fields that needed to be updated in real time were the latitude and longitude fields, and the time (and the checksum value). Time was simply calculated using the system time and then formatted to the correct string structure, and the latitude and longitude values determined by converting the Eastings and Northings of the users position in the environment to their latitude and longitude counterparts.

Once the character string that would be used for the sentence body was created, the checksum value needed to be calculated and added (Figure 28). This value is a two digit hexadecimal value (from 0 to F, with A=10 and F=15) indicating the logical XOR comparison between all the bytes making up the character string. Once the sentence had been acquired by the receiver, the same bitwise sum operation would be performed on the character string and if the two checksum values did not match the sentence would be discarded as this would show an error in the data transmission.

```
static byte GetChecksum3(char* sentence)
{
	// Loop through all chars to get a checksum
	int Checksum = 0;
	for(int i = 0; i < strlen(sentence); i++)
	{
		char Character = sentence[i];

		if (Character == '$')
		{
			// Ignore the dollar sign
		}
		else if (Character == '*')
		{
			// Stop processing before the asterisk
			break;
		}
		else
		{
			// Is this the first value for the checksum?
			if (Checksum == 0)
			{
				// Yes. Set the checksum to the value
				Checksum = ((byte)(Character));
			}
			else
			{
				// Now XOR the checksum with this
character's value
				Checksum = Checksum^(byte)(Character));
			}
		}
	}

	byte Hex = (byte)Checksum;

	return Hex;
}
```

Figure 28: Checksum code method

Once calculated, the byte is added onto the end of the sentence in hexadecimal format before being transmitted. It is important that it is in fact the byte that is being transmitted and not the textual representation, so for example if the checksum value was 8A it is important that it is in fact the hexadecimal value (equating to 10001010 = 138 in decimal) and not the ASCII characters (8 = 00111000, A = 01000001) that is added to the sentence.

Once the GPS sentence had been constructed, transmission was performed using the Bluetooth wireless network system. Using a Windows based PC this transmission can be performed by sending the data to a COMM port on the system set up as a Bluetooth port. In the program code, the COMM port was opened and set up using standard parameters (baud rate, parity bits etc.) and then to send the sentence, the string of characters was written to the COMM port using Windows file handlers. The process of writing to the COMM port was repeated every second until the system was closed down or the connection with the Bluetooth device severed. In the case that the system could not connect with the PDA device, the program re-attempted to connect until either a connection was made or the program exited.

6.2.4 Training Ground Construction

Owing to the need of user control within the environment, it was decided important to allow the test subjects to get used to the user interaction controls before they began the experimental task itself. Therefore time was allowed for the users to explore within a training environment. This environment used the same controls (motion model) as the experimental environment and the same imagery for roads and ground surfaces. The elevations, road network and land use were all artificially generated using photo editing and GIS software (Figure 29).

Firstly, elevation was generated using Adobe Photoshop by generating a greyscale cloud texture, which was then adjusted by altering contrast values and applying a Gaussian Blur to create a smoother terrain. It was purposefully made that the image used would contain a large number of peaks and troughs to ensure that the user was forced to look around more when moving within the terrain so that they would gain experience of this interaction. Exporting this texture as a TIFF image allowed for the importing into Blueberry3D where the pixel values could be used as an elevation map. Landcover classifications were then generated by creating multiple layers within Photoshop, each containing coloured regions generated by rendering clouds, increasing contrast to maximum, deleting the black areas and then colouring the remaining areas with a single colour. These layers were then blended using different blending options (i.e. colour dodge, colour burn, overlay etc.) to result in a single image comprising of multiple coloured regions which were then assigned different landcover classifications within Blueberry3D. The road network within the training environment was generated manually by drawing vector shapefiles within ArcMap which followed contours of the elevation image.

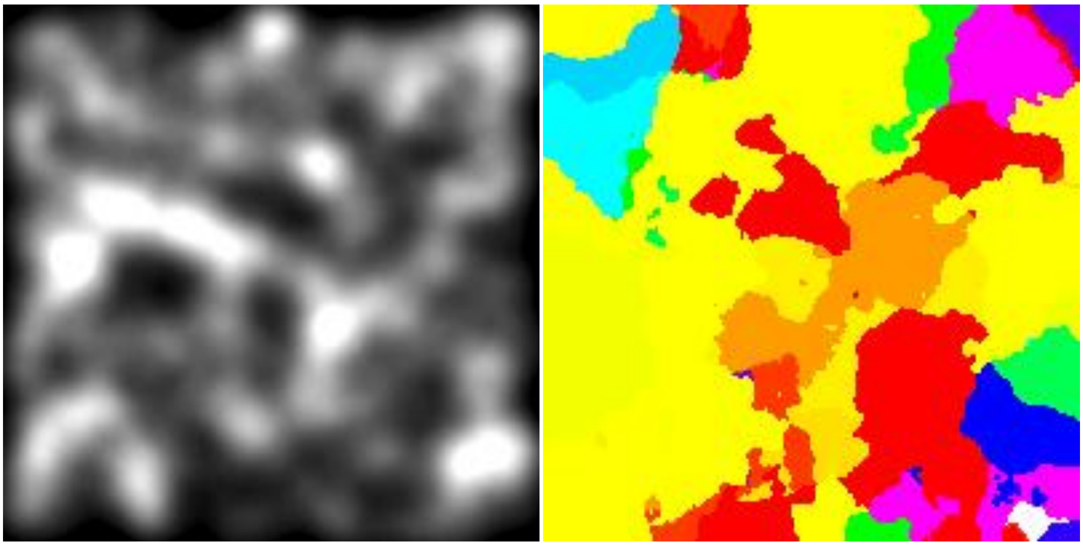


Figure 29: Training ground elevation (left) and land-use (right)

For the training ground, it was felt that rather than making the environment look realistic, it was more important to introduce a large amount of different visualisations and varying terrain heights to encourage the user to look around and move within the environment as the whole area could not be seen from a single view point. A view from within the training ground can be seen in Figure 30.

Figure 30: Training ground

6.2.5 Experimental Design

To assess the effects of the info-marks on the route tracing exercise users undertook a trial consisting of three parts to collect both quantitative and qualitative data:

1. Gaining experience of the controls in a training environment
2. Undertaking the experiment itself within the experimental environment
3. A short semi-structured interview

Before taking part in the trial, all users were asked to read and sign an informed consent form and given the opportunity to ask questions. They were also given the opportunity to decline the use of audio recording equipment and the use of quotes in produced documents. Ethical clearance was obtained from the University of Leicester Geography Department Ethics Committee before commencing the trial portion of the study.

The first section of the trial was to use the smaller artificially generated environment discussed, to allow the user to acquire the skills needed to control their movement within the environment. A separate environment was used for this training portion of the trial so that the test subject did not gain any knowledge about the spatial layout of the testing environment before beginning the route tracing task. At the beginning of the training exercise, the users were shown how to use the mouse and keyboard to interact with the environment and told to inform the researcher when they were comfortable with the controls. They were also informed that they could go through and cross any feature in the environment but they would always be clamped to the ground.

Once the user had become accustomed with the controls and felt they were ready to begin the task the environment used for the experiment was loaded. The experimental task required the test participant to attempt to travel within the VE along a route shown on a paper map (Figure 31). This map displayed information regarding the general topography of the region, rivers, roads, buildings and any power lines in the environment. Standard British road atlas symbology was employed as to make the map more familiar with a large portion of the test subjects. In addition, the route to be followed and the location of the info-marks were depicted using the names displayed on the info-mark text provided to the user. Although not all users would be presented with the info-marks, it was felt important to use the same map so that if the locations depicted were used within the trials where the info-marks were present but not as a link with the info-marks, the same methods could be employed when no info-marks were present.

The route taken and map shown was the same for all participants, although whether or not text relating to the info-marks was displayed changed between test subjects based on trial number using a rolling sequence so that each presentation method had the same number of trials where it was used. Users were instructed as to what they were required to do and informed about the interface used (i.e. the compass display). If the info-mark text was to be displayed either on screen or via the PDA, the user was informed that this would appear once they were within one kilometre distance from the corresponding point on the map.

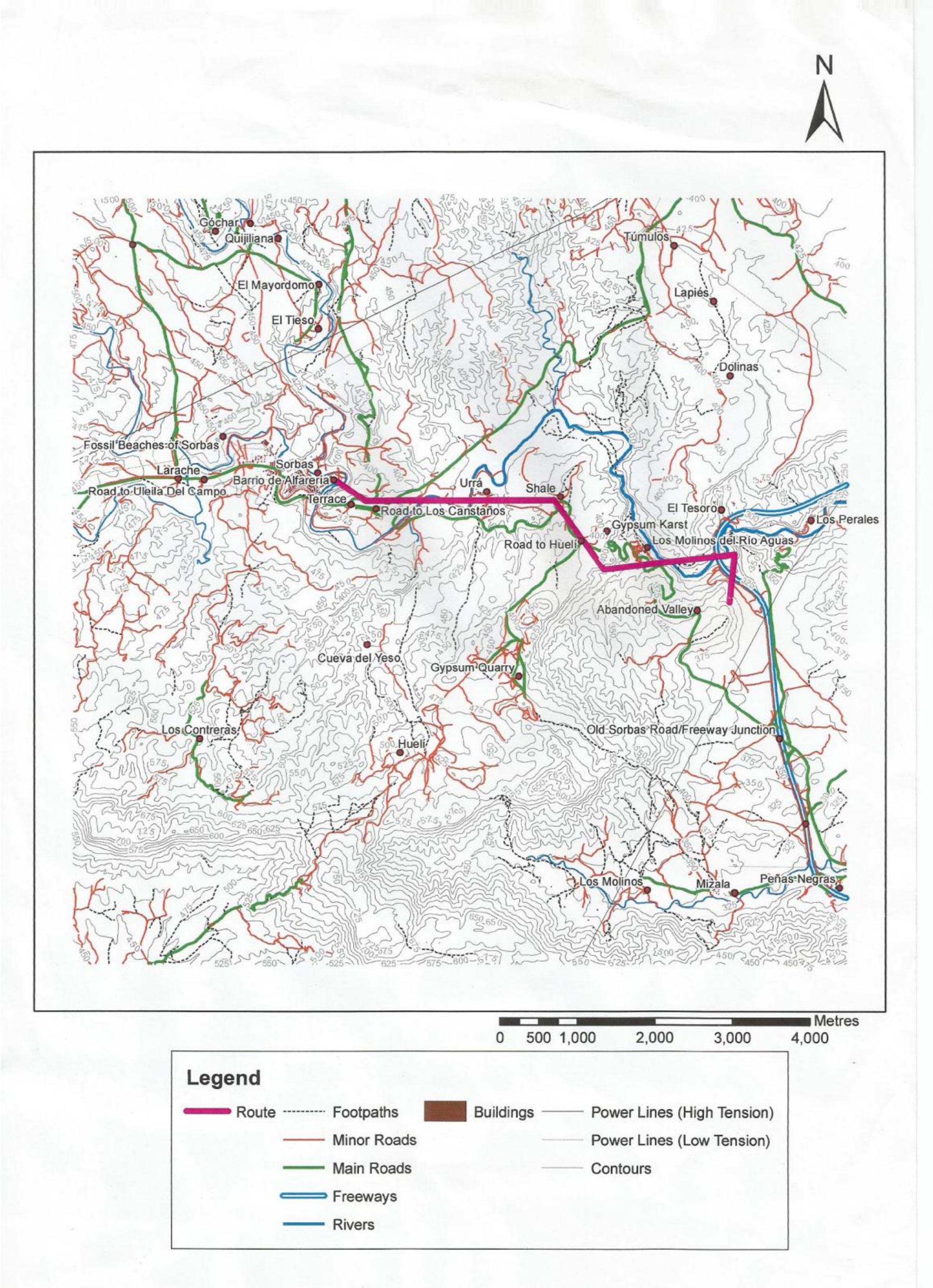


Figure 31: Map given to users

Additionally, a sheet was provided showing images of the different road types so that the user would know what type of road they were looking at in the environment, as well as being told that the river would be pointed out to them when it first enters their view as this was a moving texture which is difficult to depict on paper. The user was instructed that they could ask questions about what type a feature was and about controls, but they could not ask about where they currently were or what a specific feature was (i.e. the name of a settlement). Additionally, users were informed that the points of interest identified on the map were not necessarily settlements, and that there was no marker indicating the end point and so they needed to inform the researcher when they believed they were at the correct end location. Before commencing the task, subjects were informed that the audio device would be set recording (if agreed at the informed consent stage) and asked to verbalise their thoughts throughout the task. After instructions were given, the user was moved to the starting location and logging of movements began.

Upon indicating that the subject believed themselves to be at the correct end point, the computer system was terminated and the short interview begun. Questions asked were as follows:

- Are you male or female (as to record on the audio device for later analysis)
- From the following age bands, which one applies to you:
    - Younger than 15
    - 15 to 20
    - 21 to 30
    - 31 to 40
    - 41 to 50
    - 51 to 60
    - 61 to 70

    - Older than 70

- On a scale of one to ten, with one being rubbish and ten being really good, how would you rate:
    - Your skill with using maps
    - Your navigational skills
- In general, how easy do you think it was to travel along that route?
- On a scale of one to ten, with one being very hard and ten being very easy, how would you rate how easy it was?
- In general, how accurate do you think you were?
- On a scale of one to ten with one being not at all accurate and ten being bang on the line, how would you rate yourself?
- Are there any parts along the route that you found more difficult than others?
- Are there any parts along the route that you found easier than others?
- What did you generally use in the environment to help you complete this task?
- What on the map did you generally use to help in the task?

If info-marks were displayed in the trial, the user was also asked:

- Do you think the pieces of text helped in the task?
- Did you use them more as a means of determining where you were or where you had to go?

Questions were also asked to all participants about particular types of features within the environment:

- Did you notice any linear features in the environment that you could travel along or could direct your travel?
- Were there any features that you felt you should not or could not cross?
- Did you notice any areas that were different in appearance or had some unique characteristic that set them apart from their surrounding areas?
- Were there any locations or features that were integral parts on the task (turning points, start and end points) or were key decision making points on this or any other pathway within the environment?
- Did you notice any unique features that helped you to determine where you were or where you needed to go?
- Out of all the features mentioned, which ones did you use most?

Finally, questions were asked about the general feeling towards the task and environment, and any ways that the task could be made easier:

- Can you think of any ways that the task could be made easier, such as adding, removing or changing objects and features?
- What are your general thoughts of the task as a whole and the environment?

If any question was not understood fully by the participant then an attempt was be made to explain what is meant. Also, if an aspect of a subject's response warranted further investigation then additional questions were asked in response.

6.2.6 Data Extraction

Data was extracted for analysis from the track logs and interview transcriptions. The methods used to extract and analyse this data is discussed in the following sub-sections.

6.2.6.1 Track Logs

Information from the track logs was extracted by converting the text files into ArcMap Shapefiles. This was performed using custom VBA code that read each file and created a point feature in a shapefile for the first position log of each second. Prior to the text files being run through the process and converted to the point features, the logs were cleaned to remove log points when the user was stationary at the beginning and end of the trial. It was decided that all entries in the log file would be removed that were beyond two full seconds before the first movement and past two full seconds after the last movement. The first and last movement was seen as either rotational or locamotive. Once imported into ArcMap as a Shapefile, each point track log was also converted to a line Shapefile using inbuilt functionality.

The shapefiles created were then used to extract several metrics that could be used as a means of analysing task performance (Table 12). In this context, accuracy is used to denote that the metric identifies the level of deviation from the intended route, and efficiency signifies how efficient the user was in completing the task in terms of time needed to reach the end point and the amount of backtracking needed. A low efficiency value would indicate that the user needed to perform a large number of corrective measures or spend a large amount of time determining their location. A more efficient user would be able to pause less and require less corrective measures, even if they were not particularly accurate.

Metric	Aspect
Area	Accuracy
Distance from end point	Accuracy
Path length	Efficiency
Time taken	Efficiency
Average speed	Efficiency

Table 12: Metrics used

Each of these metrics were extracted from the GIS track dataset using a series of inbuilt functions. The path length metric was obtained using the HAWTHS Tools toolset (see Beyer (n.d.)) and the time taken determined by a simple time difference between the first point and the final point of the point format track log. Average speed was then determined using these two values. Distance from end point was again calculated using the HAWTHS toolset after extracting the end point from each line format track log using the 'Convert feature vertices to point' ArcMap tool and selecting to only extract the end point. These individual point datasets were then merged into one point shapefile along with the actual end point obtained from the intended route. Running the 'Distance between points tool (within layer)' resulted in a distance matrix being output as a .csv file which could then be interpreted to obtain the distance of each end point to the intended end point.

Finally, the area metric was calculated by merging each of the linear track logs with the intended route shapefile to obtain a dataset comprising of two lines for each trial. The start and end points of these two lines were then joined by adding additional lines to the dataset, before converting the lines to a polygon using the 'Convert lines to polygon' tool. After running the tool, the multiple polygons in each dataset were merged to form one polygon for each trial, and then the area calculated (Figure 32). The area metric can be used to give an idea of how close to the line the subject was travelling. If the user was always travelling on the line then the area value would result in 0, whereas if they deviated largely then the area value would also increase.

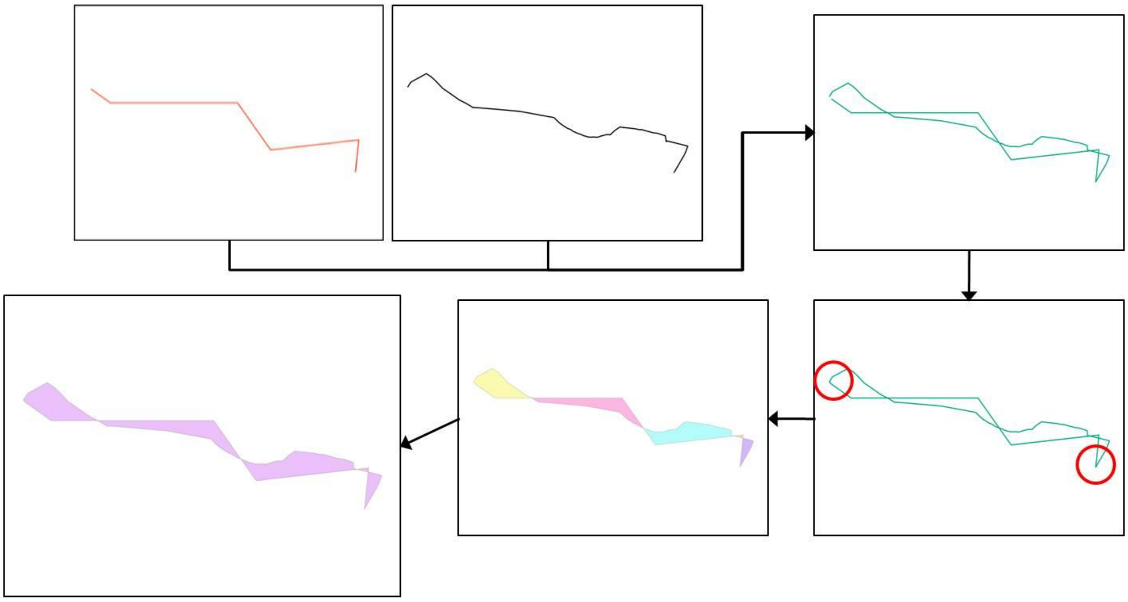


Figure 32: Calculation of area metric

Although each of the aforementioned metrics can be used to investigate specific aspects about the tracks collected, they only provide performance and efficiency indicators for that one specific attribute. However, it may be the case that one metric indicates high performance whereas another may identify that large errors occurred (Figure 33). For example, if a person made several mistakes along the route, but then retraced their steps to return back to the intended route then the area metric would be low but the overall length would be large. Additionally, a person could travel the same distance on their route as the overall length of the intended route, but actually be nowhere near the intended route and so have a large area metric. Finally, the distance from end point should also be taken into account as the person could travel perfectly along the route but massively overshoot the end point. In this instance, area would be small, the total travel distance would be larger than intended but not necessarily significant, but the distance from end point would be large.

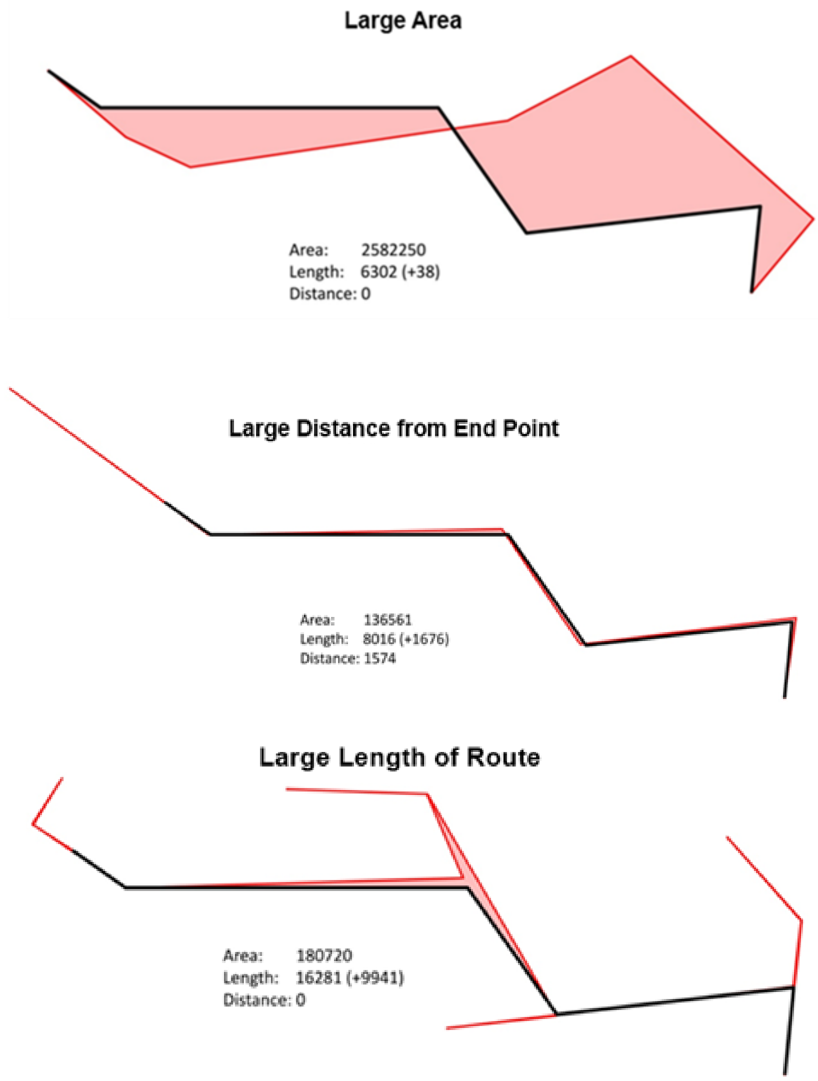


Figure 33: Extreme individual metric values

It was decided therefore to combine the three performance metrics into a single statistic that could be used for analysis of overall performance. As the numbers generated for each metric indicate different aspects of the task and so have different scales, variability and dimensionality (distance from end point and length are one dimensional linear values, although distance from end point would tend to be much smaller, and area is a two dimensional area value) relational values were needed with regards to the individual metric population.

$$m = \frac{x - \bar{x}}{\sigma_x} \tag{6.2}$$

Where m is the individual metric relational value, x is the metric value, $\bar{x}$ is the mean value of the metric population, and σ is the standard deviation of the metric population. By combining the relational value for each metric a single overall performance metric for the trial is generated.

$$m = \frac{a - \bar{a}}{\sigma_a} + \frac{l - \bar{l}}{\sigma_l} + \frac{d - \bar{d}}{\sigma_d} \tag{6.3}$$

Using this equation, it can be seen that all of the performance metrics are taken into account, so if a person did particularly well on two metrics (say length and distance from end point) but poorly on another (area) then this poor value would in turn reduce the overall performance metric. If needed, modifications could be made to allow for weighting thus giving one metric more influence on the overall combined value although this was not implemented in this study.

6.2.6.2 Interviews and In-Trial Feedback

Transcripts were made of all interviews conducted as well as the associated recordings of in-trial feedback, and analysed by extracting key information. The data extracted and analysis methods varied depending on the question type within the interviews and the aspect that was being investigated. For most responses the actual numerical or verbal response was used in analysis with the exception of the Urban Image questions and the improvements question where responses were grouped between interviews.

One of the issues arising from the responses to the Urban Image questions within the interviews were that multiple definitions were given for the same feature type, such as *"main roads"* and *"highways"* referring to the same main road features, and *"agricultural land"*, *"farms"*, *"orchards"*, *"vineyards"* and *"plantations"* all referring to the linear plantation formations within the environment. Therefore, after extracting all responses from the interviews the different definitions indicating the same feature were aggregated together. Any features purely relating to software rendering glitches (i.e. buckling of roads) were not counted as features.

After aggregation occurred, another aspect identified within the responses were that several features all shared some form of a similar characteristic which may be the reason for their identification within the classification type. Owing to the relatively small sample size and varied responses, some classifications did not have an apparent 'leader' as many features were mentioned only a few times each. Therefore, to get a more general understanding of the types of feature, features with similar characteristics were grouped together into 'super-classifications'. Care was needed however not to join a feature into one of these super-classifications if on its own it stood out as a key feature in the classification group, or that two aspects of the same feature were not joined if the aspects themselves were different (i.e. road direction and bends in a road). In some instances however, if a feature that stood out on its own was a general descriptor of other responses (i.e. roads is a general descriptor of main roads, minor roads and freeways) then the specific features were grouped into the general descriptor. Specific mentions about a particular aspect of a feature were not grouped in with the general feature as the mention for the general feature could imply multiple factors and not just the one that was explicitly identified by other users. The final list of features and super-classifications are referred to as end-feature types

The ‘number of people count’ was used to identify the number of test subjects who indicated the end-feature type, having mentioned it at least once in the interview question response. For example, if a person gave a response of *“Roads, minor roads and highways”* for a question, then if all three features mentioned were part of a single super-classified end-feature type, say Roads, a count value of one would be recorded as opposed to three. This is because although three particular features were mentioned only one end-feature type was (three instances of the Road end-feature type). By collating the data in this way the percentage of users who mentioned a particular end-feature and so identify the ‘hardness’ of that feature as used by Omer et al. (2005) can be assessed. If a feature is mentioned by a large percentage of interviewees then it is seen as hard in that the feature would occur in most people’s Urban Image. On the other hand, if it is only mentioned by a small percentage then it would be seen as soft in that only a few people would identify that particular feature in their Urban Image. A particularly hard feature would have a percentage of users close to 100% whereas soft features would approach 0%.

Similarly to the Urban Image responses, the feedback from the interviews on possible improvements to the environment were also grouped into particular types. Initial grouping was again based on responses that imply the same aspect such as *"Road signs at junctions"* and *"Sign posts on junctions"* and then further grouping was used to split responses into a general improvement aspect such as *"Road signs at junctions"* and *"Sign posts at route crossing a road or river"* being marked as directional highlighting improvements.

6.2.7 Data Analysis Methods

In this study, data were collected and analysed in both quantitative and qualitative forms. The methods employed in the analysis of the data are now described.

6.2.7.1 Quantitative Data

For statistical testing to be performed on the data collected from the track logs, the shape of the distribution needed to be determined. The data collected for each metric was initially assumed to be similar to a normal distribution which was tested using the Kolmogrov-Smirnov test (KS-test). To do this, the mean, standard deviation and sample size was recorded for each treatment group in each metric and the KS-test used to compare the actual data with a sample of the same size from a normal distribution using the same mean and standard deviation values. If the p-value obtained was greater than 0.05 then it was decided that the actual data was of a normal distribution as there would be no significant difference between the distribution of the actual data and a true t-distribution with the same mean and standard deviation.

However, owing to the smaller sample number (10 for each treatment group) a normal distribution cannot be used as a description for the sample and so instead a t-distribution is used. A t-distribution is similar to a normal distribution in its shape and characteristics, but it is less peaked around the mean and more dispersed, with the dispersion varying according to the sample size (Rowntree 2003). However, as the sample size increases and reaches 120, the t-distribution becomes practically identical to a normal distribution (Diamantopoulos & Schlegelmilch 1997). All tests performed were two-tailed as whether there was any effect by presenting the info-marks was to be assessed, not just whether it improved or made worse performance.

In all testing, the hypotheses were:

- H_0: There is no difference in the metric between the treatment groups ($\mu_1 = \mu_2$)
- H_1: There is a difference between the treatment groups for the metric ($\mu_1 \neq \mu_2$)

As t-tests were used, only two samples can be compared at a time, so the tests for each metric were:

- No info-mark display vs. on-screen info-mark display;
- No info-mark display vs. PDA info-mark display;
- On-screen display vs. PDA info-mark display.

Therefore the hypotheses were tested between each treatment group by running the t-test multiple times within each metric.

The test statistic used in the t-test is calculated the same way as a z-value in that it is the standard error of differences between the sample means (Rowntree 2003). The actual method for determining the t-statistic depends on whether the variances of the two groups can be seen as the same and whether or not the groups contain the same number of observations. If the variances can be seen as equal and both groups have the same number of observations then the following equation is used to calculate the t-statistic:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1X_2} \times \sqrt{\frac{2}{n}}}$$

6.4

Where

$$S_{X_1X_2} = \sqrt{\frac{1}{2}\left(S^2{}_{X_1} + S^2{}_{X_2}\right)}$$

6.5

If it is the case that the variances can be seen as equal but there are not the same amount of observations in each group, the t-statistic is calculated as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1X_2} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6.6

Where

$$S_{X_1X_2} = \sqrt{\frac{(n_1 - 1)S^2{}_{X_1} + (n_2 - 1)S^2{}_{X_2}}{n_1 + n_2 - 2}}$$

6.7

Finally, if both the number of observations and the variances differ between the two groups, then the t-statistic is calculated as:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$$

6.8

Where

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s^2{}_1}{n_1} + \frac{s^2{}_2}{n_2}}$$

6.9

To evaluate the significance of the t-statistics, the degrees of freedom value (df-value) also needs to be calculated for each test. Again, the method for calculating this value depends on whether the variances and observation counts are equal in both groups. If they are both equal, the df-value is calculated with:

$$df = 2n - 2$$

6.10

If the number of observations differ but the variances can be the same then the following is used:

$$df = (n_1 + n_2) - 2$$

6.11

Finally, if the number of observations and variances cannot be seen as equal, then the df-value is derived using:

$$df = \frac{(s^2{}_1/n_1 + s^2{}_2/n_2)^2}{(s^2{}_1/n_1)^2/(n_1 - 1) + (s^2{}_2/n_2)^2/(n_2 - 1)}$$

6.12

In all the tests conducted, the observation count was always the same for each group (n=10) but the variances needed to be tested for equality. Preliminary investigation of this was performed using the Bartlett test of homogeneity of variances for each metric. This test analyses multiple samples to determine whether the variances can be seen as equal with each other. If this was identified to be the case for the set of treatments in a metric, then an f-test was performed between each treatment group in that metric to determine which groups had different variances. The t-tests were then performed depending on the results of these variance tests. The reasoning behind performing the analysis of variances before performing the t-test was to identify which type of test would be performed by the R statistical package. By default, R will select which t-test (Student's or Welch's) based on the data provided, but it was felt important to gain an understanding of why a particular test was performed rather than relying on R to make the correct decision without question. As the tests for equal variance could be conducted very easily within R using the already entered data, their use appeared prudent.

6.2.7.2 Qualitative Data

Qualitative data in the form of interview responses and in-trial feedback was analysed in different ways depending on the information that was to be extracted. The numerical scaling of information regarding accuracy and ease was statistically analysed to evaluate how the presence of info-marks affected the user's perception of these factors. Other data was analysed by extracting the key aspects of responses again depending on the question being asked. The two areas that the data were collected each had a particular role in the understanding of the performance during trials. The interviews allowed for a more in-depth probing of performance and reasoning behind decisions made, whereas the in-trial feedback allowed for the collection of more immediate expressions of emotion and decision making factors. By collecting the data using these two methods it is possible to understand performance without entirely relying on the users memory, whilst allowing for in-depth investigation of key aspects.

After the interviews were conducted the audio recordings (containing both the interviews and in trial feedback) were transcribed for analysis. Each transcript was then analysed by extracting the relevant information from the text and then counting the occurrences between transcripts. The general words and phrases that were selected from the interviews revolved around features within the environment, specific navigation methods and general feelings towards the task. During the in-trial feedback data extraction attention was focussed on similar elements although signs of confusion or happiness were also taken into account (i.e. positive exclamations when info-marks appeared).

As the data collected from the rating questions (perceived accuracy and ease of task) was ordinal in nature as opposed to continuous, Mann-Whitney tests were implemented in analysis as the data was ordinal and non-paired. If a response given was between two values (i.e. *"I'd say a seven or an eight"*) then the lowest of the values was used to ensure that only the whole number values were used in the analysis. Although the inclusion of decimal numbers would not have an effect on the statistical test used, it was felt better to keep the responses that were to be used in analysis on the same whole number scale. This meant that a number could be directly matched to a whole number on the scale as opposed to being rounded up or down at the analysis stage.

6.3 Results

30 participants (10 in each delivery method group) took part in the user trials consisting of 12 male and 18 females with an average age of 29 based on age bands. Participants were predominantly students or staff within the geography department of the University of Leicester. Each test subject was asked to rate their skill with maps and navigation on a scale of one to ten with one being awful and ten being excellent. Results for the skill with maps rating were a mean value of 6.97 and a standard deviation of 1.45, and results for the skill with navigation question were a mean of 6.37 and a standard deviation of 1.85, indicating an average to above average skillset for participants in both aspects. All statistical analysis was performed using the R statistical analysis software. Some participants had also taken part in the trials for the first experiment of this thesis, but this was deemed not to cause any problems due to the difference in task and the fact that it would be very difficult to identify position within the navigation tasks from any information gained from the first distance estimation investigation. In addition, due to the minor changes in the appearance of vegetation in the environment it was felt that such transfer of information (if initially possible) would be far more difficult.

6.3.1 Quantitative

The first stage of data analysis was to determine what type of distribution each of the treatment groups in each metric were. This was performed using a Kolgorov-Smirnov test against each group and a normal distribution of the same parameters which can be seen in Table 13.

	Area	**Length**	**Distance**	**Combined**	**Duration**	**Speed**
None						
Min	391738.5	226.9297	29.1173	-2.73538	563	3.428231
Max	3624558	4419.057	2985.53	8.524557	1974	7.659839
Mean	1279047	1317.552	566.7073	0.240417	1283.1	5.844603
SD	1019232	1355.591	900.7383	3.433075	397.3925	1.175863
Median	1025338	679.0446	258.0339	-1.18952	1284	6.030466
On Screen						
Min	370691.8	376.3788	42.32254	-2.31475	804	4.574624
Max	2563801	3728.74	465.0956	3.062803	2013	11.94709
Mean	1168536	1772.327	196.6794	-0.1468	1315.9	6.676225
SD	725312.4	1201.207	163.8182	1.648741	431.6275	2.202596
Median	1021019	1617.088	96.00522	-0.33789	1201.5	6.334595
PDA						
Min	589523	55.40406	59.06039	-2.04884	713	4.753533
Max	2680715	8776.056	1255.156	7.549828	2407	8.969817
Mean	1236381	1415.398	355.8074	-0.09361	1115.9	7.089421
SD	744026.1	2623.791	353.3158	3.117879	477.353	1.332448
Median	933284.5	531.7743	247.4601	-1.53646	981	7.075809

Table 13: Track metric distributions

6.3.1.1 KS-Test Results

Kolgoromov-Smirnov tests were performed on each treatment within each metric group to identify normality. Test results (Table 14) for all groups do not identify a significant difference (p-value of less than 0.05) between the data collected and a normal distribution of the same parameters, and so therefore it can be assumed that all of the treatment distributions can be seen as a normal distribution.

	Area	**Length**	**Distance**	**Combined**	**Duration**	**Speed**
None	0.787	0.4175	0.1678	0.1678	0.9945	0.9945
On screen	0.4175	0.9945	0.787	0.4175	0.787	0.9945
PDA	0.4175	0.4175	0.9945	0.05245	0.1678	0.9945

Table 14: KS test results

6.3.1.2 Tests for Equal Variance

The first stage for testing for equal variance between treatment groups within each metric was to perform a Bartlett test on the data for each metric. The results from these tests identified which form of t-test would be performed between groups. Results (Table 15) indicated that a significant difference between variance could be found in the track length and distance from end point metrics (p=0.03899 and p=2.219e-5 respectively). This indicates that between the treatments (none, on-screen and PDA display methods) for these groups there was a significant difference in variance. Therefore the next stage was to identify between which treatments these differences occurred, undertaken using an f-test between the treatment groups in the metrics with different variances.

Metric	Bartlett test p-value
Area	0.5215
Track length error	0.03899
Distance from end point	2.219e-5
Combined performance metric	0.1052
Travel duration	0.865
Average travel speed	0.1340

Table 15: Bartlett test results

From the f-test results (Table 16) it is clear that a significant difference in variances can be found within the track length error metric between the onscreen and PDA delivery methods, and within the distance from end point metric all treatments have significant differences in their variances.

Metric	None vs. screen	None vs. PDA	Screen vs. PDA
Track length error	0.7245	0.06219	0.0293
Distance from end point	2.27e-5	0.01023	0.03173

Table 16: F-test results

6.3.1.3 T-Test Results

Depending on the results from the equality of variance tests, either a Students t-test or a Welchs t-test was performed between the treatments in each metric. If the variances were found to be different then a Welchs t-test was performed, else a Students t-test was undertaken for analysis.

Area: The area metric can be used as identifying the general deviation from the intended route, with a larger value indicating a larger distance from the intended route. As the variance tests indicated that the variances were equal between all treatments within the area metric then a Students t-test was performed between each treatment. Results from the t-test can be seen in Table 17, with a boxplot of the datat shown in Figure 34.

	Df	t-value	p-value
None vs. onscreen	18	0.2794	0.7832
None vs. PDA	18	0.1069	0.916
Onscreen vs. PDA	18	-0.2065	0.8387

Table 17: Area metric t-test results

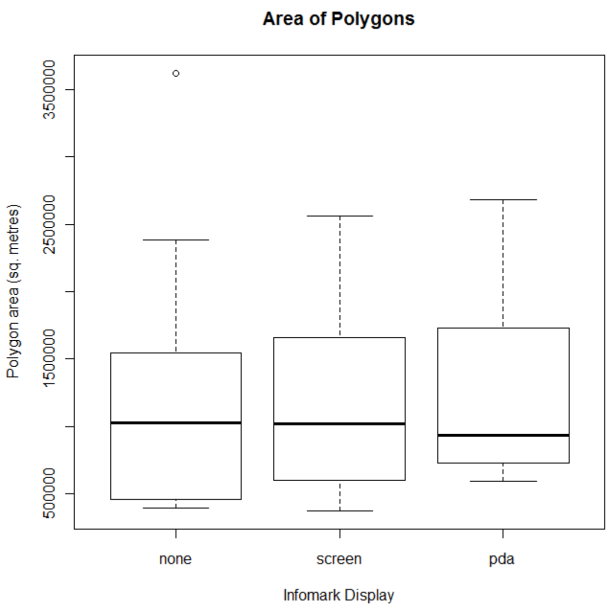


Figure 34: Area metric boxplot

From these results, it can be concluded that there is insufficient evidence of a difference between the treatment groups, and so the presentation of info-marks had no effect on the level of deviation from the intended route.

Track Length: For the track length error metric, results from the Bartlett and f-tests indicated that there was a difference in variance between the onscreen and PDA delivery methods. Therefore, Students t-tests were applied between the none and onscreen, and the none and PDA treatments with a Welch t-test applied to the onscreen vs. PDA treatments. Results from the tests for the track length metric can be seen in Table 18 with a boxplot depicted in Figure 35.

	Df	t-value	p-value
None vs. onscreen	18	-0.1048	0.9177
None vs. PDA	18	-0.794	0.4375
Onscreen vs. PDA	12.614	0.3911	0.7022

Table 18: Track length metric t-test results

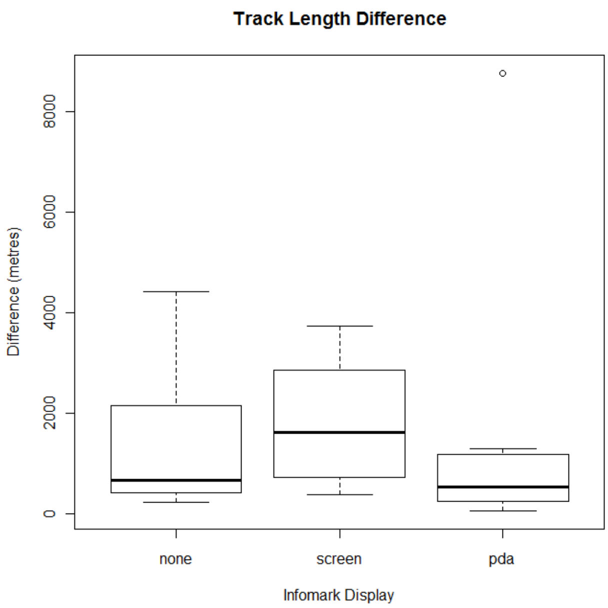


Figure 35: Track length metric boxplot

These results indicate that there is insufficient evidence to reject the null hypothesis that there is no difference between the treatment groups.

Distance From End Point: When analysing the distance from end point metric, the Bartlett test again indicated that variances between treatments in the distance from end point metric were significantly different, with further f-tests identifying that variances between all treatment groups within this metric were significantly different. Therefore, Welch t-tests were used to compare all treatments with the results being seen in Table 19. A boxplot of the data is shown in Figure 36.

	Df	t-value	p-value
None vs. onscreen	9.595	1.2781	0.2313
None vs. PDA	11.705	0.6893	0.5041
Onscreen vs. PDA	12.699	-1.2921	0.2193

Table 19: Distance from end point metric t-test results

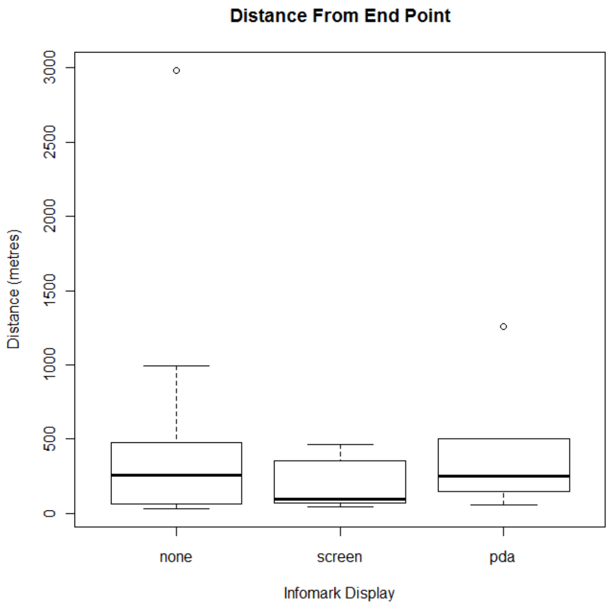


Figure 36: Distance from end point boxplot

In this metric it is also the case that there is insufficient evidence to reject the null hypothesis of no difference between treatment distributions.

Combined: The equation for combining the performance metrics (area, track length error and distance from end point; see section 6.2.6.1) was used to create a single metric used to indicate overall performance in the task. Tests to determine equality of variance indicated no significant difference between the treatment variances so Students t-tests were applied for analysis with the corresponding results being seen in Table 20, with Figure 37 showing the data in a boxplot.

	Df	t-value	p-value
None vs. onscreen	18	0.3215	0.7515
None vs. PDA	18	0.2278	0.8224
Onscreen vs. PDA	18	-0.0477	0.9625

Table 20: Combined metric t-test results

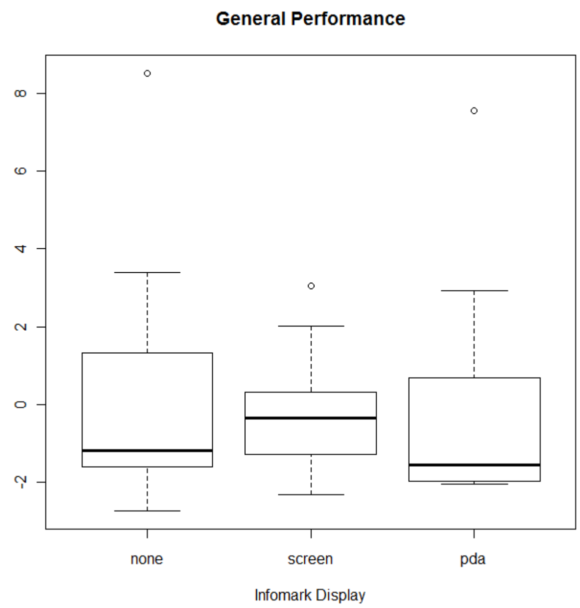


Figure 37: Combined metric boxplot

Again, no significant values have been identified between treatments, so the null hypothesis of no difference between treatment groups is accepted.

Travel Duration: The travel duration metric is the time taken to reach the believed end point. This data was collected using the time stamps on the track logs and was analysed using the number of seconds. Variances could be seen as equal and so a Students t-test was conducted between display methods. Values from the tests can be seen in Table 21 and a boxplot seen in Figure 38.

	Df	t-value	p-value
None vs. onscreen	18	-0.1768	0.8616
None vs. PDA	18	0.8513	0.4058
Onscreen vs. PDA	18	0.9827	0.3388

Table 21: Travel duration metric t-test results

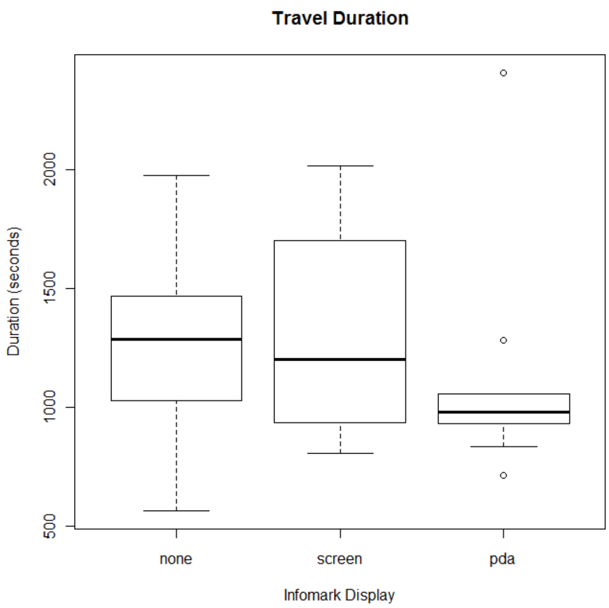


Figure 38: Travel duration metric boxplot

Results indicate no significant difference between display methods with regards to total time taken, although several outliers can be seen in the PDA dataset.

Average Speed: Average speed was calculated by dividing distance travelled by total time taken. Again variances were tested using f-tests and shown not to be significantly different and so Students t-tests were performed between display types, results being found in Table 22. Figure 39 shows the data in a boxplot representation.

	Df	t-value	p-value
None vs. onscreen	18	-1.0533	0.3062
None vs. PDA	18	-2.2151	0.03989
Onscreen vs. PDA	18	-0.5076	0.6179

Table 22: Average speed metric t-test results

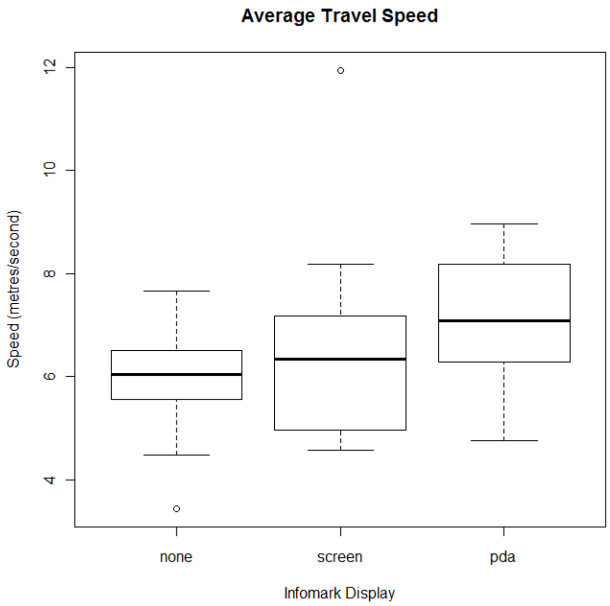


Figure 39: Average speed metric boxplot

Results from the two-tailed Student t-tests indicate that although there is no significant difference between no info-mark display and onscreen display, and onscreen and PDA display there is a significant difference between no info-mark display and when info-marks are displayed on the PDA (p=0.03989). This result means that there is evidence showing that the average travel speed is faster when info-marks are displayed on a PDA as opposed to when they are displayed on-screen.

6.3.2 Qualitative

Qualitative data was collected via verbal feedback through the post-trial interviews and in-trial reporting.

6.3.2.1 In-Trial Reporting

During the trials people were asked to verbalise their thoughts whilst navigating the environment. Quality of results was mixed owing to a number of users not providing feedback and some speaking but in a way that was difficult to pick up by the recording device.

When analysing the data, the transcripts produced from the recordings were investigated by identifying any mention of a particular feature such as a road, river or topological feature. Also extracted were mentions of compass direction, info-marks and building structures. If a feature was mention multiple times each mention would be counted as this could highlight the importance of that particular feature. For example in one trial the following was recorded:

> *"Go along the road to check where I think I am. So the road should bend around to the right in a moment according to the map, and to the left. So I'm looking for the hairpin in the road so I can then turn due west."*

In this block, the road is mentioned three times (*"... along the road to check... the road should bend... looking for the hairpin in the road"*) and so three mentions would be recorded although the user is speaking about the same road. It is important to record all three of these as three distinctive uses can be seen:

1. Travelling along the road
2. Expecting a particular aspect of the road to confirm location
3. Using a specific feature in the road to identify a turning point

Another example would be

> "*I need to head sort of towards the freeway, and then cross over the freeway"*

Here again two recordings for road would be made: one as a directional cue, and another as identifying that the road needs to be crossed.

With regards to compass direction there were three predominant methods that were mentioned by users. These were to describe the direction in which the user was heading or needed to be heading (i.e. *"I've got to go a fraction north"*), the position of a feature relative to the users location (i.e. *"I've just crossed the footpath and I'm on the north side of the footpath"*), and the trending direction of linear features (i.e. *"So we've got a solely south-easterly trending river"*).

One common aspect was noted in all trials where feedback was given – the use of roads. Throughout the trials users spoke about expecting to see roads and using intersection (either of their track or between roads) to identify their location in the environment. To a lesser extent rivers were also used in the same manner. As well as intersections between linear features being a predominant cue for orientation, the shape of features was also used. For example one user stated *"So just north is the bend* *in the river, so we should actually be crossing just south of that bend"* indicating that in that instance it was the bend in the river that was the main aspect used to identify location.

When the info-marks were presented to the user, often responses would be given directly after a change in the info-mark being displayed. In some instances the response was a positive response such as "*yes, that's a good sign"* and in others the response indicated the use of the info-mark as a marker to identify a turning point such as *"Shale… Ok, that means I need to turn in a minute"*. A response was recorded if the info-mark was mentioned by name or a positive/negative response made directly after its appearance.

Also, in some of the trials, users stated that they needed to head towards the top of hills to look around. Once at these locations, the method of identifying roads within the scene was used as a means of orientation.

The counts for each item can be seen in Table 23. To analyse the data however percentages were taken of the number of mentions of each type per person (for example a person mentioning roads four times and compass bearings once would have a percentage of 80 for roads and 20 for compass) as can be seen in Table 23. Values of 0% are not displayed in the table. These percentages were then used to identify which features were highlighted more as some people talked considerably more than others and so without using the percentages their opinion would have a massive influence on that of the group. By using the percentages it is ensured that each user's feedback holds equal weighting.

No Info-marks							
	Roads	Rivers	Compass	Info-marks	Buildings	Terrain	Total
	60.00	40.00	-	-	-	-	5
	36.92	18.46	24.62	3.08	4.62	12.31	65
	75.00	25.00	-	-	-	-	4
	61.54	3.85	7.69	-	11.54	15.38	26
	43.75	25.00	18.75	-	6.25	6.25	16
	50.00	50.00	-	-	-	-	2
	23.68	34.21	15.79	-	10.53	15.79	38
	39.53	25.58	26.74	2.33	1.16	4.65	86
	-	50.00	-	-	50.00	-	2
	-	40.00	40.00	-	20.00	-	5
Total	390.43	312.10	133.59	5.40	104.09	54.38	249
%age	39.04	31.21	13.36	0.54	10.41	5.44	
On-screen Info-marks							
	Roads	**Rivers**	**Compass**	**Info-marks**	**Buildings**	**Terrain**	**Total**
	27.27	-	36.36	27.27	-	9.09	22
	41.46	14.63	9.76	26.83	-	7.32	41
	52.63	26.32	21.05	-	-	-	19
	35.14	21.62	21.62	16.22	-	5.41	37
	33.33	23.33	6.67	30.00	-	6.67	30
	100.00	-	-	-	-	-	2

	50.00	-	50.00	-	-	-	2
Total	339.84	85.90	145.46	100.32	-	28.48	153
%age	48.55	12.27	20.78	14.33	-	4.07	
PDA Info-marks							
	Roads	**Rivers**	**Compass**	**Info-marks**	**Buildings**	**Terrain**	**Total**
	27.42	8.06	37.10	20.97	6.45	-	62
	33.33	20.83	8.33	29.17	4.17	4.17	24
	35.00	15.00	-	35.00	15.00	-	20
	75.00	-	-	25.00	-	-	4
	16.67	-	-	66.67	-	16.67	6
	42.86	14.29	28.57	14.29	-	-	7
	35.00	15.00	20.00	25.00	-	5.00	20
	-	-	-	100.00	-	-	1
	60.00	20.00	10.00	10.00	-	-	10
	27.78	-	55.56	16.67	-	-	18
Total	353.05	93.18	159.56	342.75	25.62	25.83	172
%age	35.31	9.32	15.96	34.28	2.56	2.58	

Table 23 : In-trial feedback (percentages)

By looking at the percentages it can be seen that for the trials where no info-marks were presented to the user the linear road and river features were mentioned heavily (39% and 31% respectively). Compass bearings were also used (13%) and in two instances the info-marks were mentioned even though they were not presented to the users. These info-mark occurrences however were where the users assumed that the point-of-interest shown on the map corresponded to a settlement that was in their visual field within the environment even though they were informed at the start of the trial that this would not necessarily be the case.

When info-marks were presented (both on-screen and via the PDA) there is still seen a high usage of roads (49% and 35%) but a drastic drop in rivers (12% and 9%). Unsurprisingly there is an increase in the use of info-marks but the usage is more predominant in the PDA interface (34%) in comparison to the on-screen display (14%). Compass bearings are still used quite heavily in both info-mark display methods (21% for on-screen and 16% for PDA) with buildings and the terrain being used scarcely (buildings: 0% on-screen, 3% PDA; terrain: 4% on-screen, 3% PDA). Table 24 shows the ranks of features mentioned for each info-mark display method.

	None	**On-screen**	**PDA**
1st	Roads (39%)	Roads (49%)	Roads (35%)
2nd	Rivers (31%)	Compass bearings (21%)	Info-marks (34%)
3rd	Compass bearings (13%)	Info-marks (14%)	Compass bearings (16%)
4th	Buildings (10%)	Rivers (12%)	Rivers (9%)
5th	Terrain (5%)	Terrain (4%)	Terrain (3%)
6th	Info-marks (1%)	Buildings (0%)	Buildings (3%)

Table 24: In-trial feedback classifications

6.3.2.2 Perceived Difficulty and Accuracy

Within the interviews, each person was asked to rate the perceived ease of the task on a scale of one to ten with one being very difficult and ten being very easy, as well as being asked to rate their own perceived accuracy at following the route again on a scale of one to ten with one being not at all accurate and ten being extremely accurate. By extracting these values from the interviews and collating them according to trial type, analysis could be performed to test the effects of the info-marks on the user's own perceived performance as opposed to actual performance.

The data for this however is in an ordinal form rather than the continuous data collected from the track logs as a response is a whole number between one and ten. With this scale, one indicated that the task was very difficult and the user felt they were very inaccurate, and ten indicated that the task was very easy and they felt they were very accurate. If a response given was along the lines of *"I'd give it a seven or an eight"* then the lowest whole number value was used for analysis. If a number such as six and a half was given, the whole number part was used so a value of six would be used in the data analysis. The scores given in the interviews can be seen in Table 25.

Ease			Accuracy		
PDA	Screen	None	PDA	Screen	None
8	5	6	4	3	5
7	8	7	7	6	5
4	6	6	6	7	3
8	4	5	7	6	7
7	4	3	7	6	7
5	6	5	8	4	3
7	6	2	6	7	2
5	5	2	8	6	6
4	4	3	7	5	3
	8	3	6	6	4
1=very difficult/highly inaccurate and 10=very easy/highly accurate					

Table 25: Ease and accuracy scoring

As the data collected is ordinal and not paired, the Mann-Whitney test was be used (Clegg 1982). In this statistical test, two treatments are compared at a time meaning that a total of six tests were required to complete the analysis. Additionally, one value was missing from the interviews (the ease of task value for trial number 3) owing to an error in the interview process, but this does not cause any problems for the Mann-Whitney test. F-tests were first conducted however to check for equality of variance between treatments as this is a pre-requisite for the Mann-Whitney test, with all results indicating equality (see Table 26).

	None vs. Screen	None vs. PDA	Screen vs. PDA
Ease of task	0.5881	0.7555	0.8307
Accuracy	0.3237	0.231	0.8274

Table 26: F- test results (p-values) for ease and accuracy responses

To perform the Mann-Whitney test in the R statistical analysis package, a Wilcoxon Ranked Sum test is performed which automatically performs a Mann Whitney test when two number lists are passed. All tests performed were two tailed and significance evaluated at the 5% confidence level.

When performing the Mann-Whitney test between the treatments for the ease of task question responses, the results obtained can be seen in Table 27.

	w-value	p-value
None vs. screen	84	0.1156
None vs. PDA	74.5	0.03875
Screen vs. PDA	92	0.5315

Table 27: Ease of task Mann-Whitney test results

Results from the statistical tests indicate that when info-marks are displayed on a PDA, users feel that the task is easier compared to when no info-marks are displayed (p=0.03875). In the other instances of none versus onscreen and onscreen versus PDA display, no significant differences were found (p=0.1156 and p=0.5315 respectively).

For the accuracy of performance Mann-Whitney tests were performed between metrics in the responses to the accuracy of following the route question, results of which can be seen in Table 28.

	w-value	p-value
None vs. screen	86.5	0.1638
None vs. PDA	72	0.01220
Screen vs. PDA	81	0.06243

Table 28: Accuracy Mann-Whitney test results

As with the ease of task responses, there is statistical evidence that people felt they were more accurate at tracing the route when info-marks were displayed on the PDA then when they were not displayed at all (p=0.01220). Again however, no significance was found between the other comparisons of no display versus onscreen display, and onscreen display versus PDA display (p=0.1638 and p=0.06243 respectively).

6.3.2.3 Use of Info-Marks

All users who had the info-marks displayed either on the PDA or onscreen were asked whether they found them useful and how they were useful. Out of all the responses (n=18) only 1 user stated that they were not useful, with the other 17 users indicating that they were useful. The user who said that they did not use the info-marks did however say that although they did not rely on them they thought they were useful to have as they showed they were around the correct area.

With regards to the use of the info-marks, a number of responses were identified. These responses were: the info-marks being used as a means of determining current location; them being used for determining whether the user is going in the right direction; or the info-marks being used for a mixture of both. In some instances, the user would respond with them being used for both but more heavily on either the locational or correct direction aspect, in which case a response was recorded for both the individual and the ‘both’ aspects. Overall 9 users stated that they used the info-marks more as a confirmation of taking the correct route, 5 identified using them to determine their current location and 9 used them for both aspects. Out of the subjects who stated using them for both, 3 thought that the identification of direction may have featured a bit more, and 2 that the confirmation of correct route was used more. 2 users did not specify which was used more. Table 29 identifies the dominant usage of the info-marks for each user who was presented them.

Trial	Confirmation of Route	Both	Identification of Location
2			
3	x		
5		x	
6		x	
8			
9		x	x
11	x	x	
12			x
14	x		
15		x	
17	x		
18	x		
20		x	x
21	x		
23	x	x	

24		x	
26	x		
27		x	x
29	x		
30			X

Table 29: Use of info-marks

6.3.2.4 Identification of UIT Features

During the interviews, test subjects were asked if they noticed any features that could fit in with the feature types of the UIT. Participants were not asked questions such as *"Did you notice any edges?"* as without knowing the definition of an edge in the theory it would be difficult to get meaningful responses. Instead, the questions asked were based on characteristics of the particular feature type, such as *"Did you notice any features that you felt you should not, or could not cross?"*. Care was given as not to give examples of what types of features fall into these categories so that the response was not influenced, unless the subject was particularly struggling to understand what was meant by the question being asked. Table 30 shows the aggregated responses of users for each feature type. Within this table, the people column details the number of people who identified the particular element and the percentage of people shows what percentage of all people identifying the feature. The percentage of occurrences column shows the portion of all feature mentions for the particular classification that the feature in hand makes up. For example, if there were 100 mentions of features for the pathway classification, and 35 of these were roads, then the percentage of occurrences value would be 35. At the bottom of each classification a general 'other' row is provided which describes the number of other features that were mentioned but did not make the top four or five features.

Feature	People	%age of People	%age of Occurrences
Pathways			
Roads	24	80.00	32.88
Footpaths	15	50.00	20.55
Rivers	13	43.33	17.81
Valleys	5	16.67	6.85
Other	*16*	*classes = 16* *mean = 1* *SD = 0*	*21.92*
Edges			
Steep Terrain	11	36.67	27.50
Rivers	10	33.33	25.00
Boundary Features	6	20.00	15.00
Buildings	5	16.67	12.50
Other	*8*	*classes = 6* *mean = 1.33* *SD = 0.82*	*20.00*
Districts			
Urbanisations	10	33.33	20.00
Plantations	9	30.00	18.00
Wooded Areas	4	13.33	8.00

Bare Ground	3	10.00	6.00
Scrubland	3	10.00	6.00
Other	*21*	*classes = 17* *mean = 1.24* *SD = 0.44*	*42.00*
Nodes			
River Crossings	9	30.00	12.33
Road Junctions	8	26.67	10.96
Roads	6	20.00	8.22
Road Bends	6	20.00	8.22
Other	*44*	*classes = 24* *mean = 1.83* *SD = 1.52*	*60.27*
Landmarks			
Roads	10	33.33	15.38
Rivers	9	30.00	13.85
Info-marks	7	23.33	10.77
Buildings	4	13.33	6.15
Road Junctions	4	13.33	6.15
Other	*31*	*classes = 20* *mean = 1.55* *SD = 0.76*	*47.69*

Table 30: "Rural Image Theory" responses

From the top four (or five if a tie occurred for fourth place) features identified based on the percentage of interviewees who mentioned them, the key features can be identified which are available in the current environment that could take the place of the more customary feature types belonging to the UIT classifications. The 'Other' feature identified in each category represents a sum of all other features identified from the interviews which did not fall in the top four (or five) features. In many of the classifications the 'Other' group is particularly high (accounting for up to 60% of occurrences in the case of the Nodes classification). The reason for these high values can be accounted to varied responses in features that can be classified which by themselves do not have enough responses to make the top four of the classification.

Pathways: The features identified in the pathway classification were the 'hardest' of features, with the term 'hardest' meaning that a large proportion of test subjects stated the same features (a soft feature would have a small proportion of test subjects identifying it). 'Roads as pathways' was the highest overall for hardness with 80% of all test subjects identifying the feature. Mentions of roads in general and specific road types were merged together as it was often the case that when one specific road type was mentioned, others were as well. Footpaths (50%) were kept separate as there was a clear differentiation between the two with them being stated separately such as

> *"I saw footpaths like in the mountains, well the hills, and the roads that you could..."*
>
> *"I did notice that there were paths and I guess the roads to walk along"*
>
> *"Roads, rivers, small pathways"*
>
> *"I would say I would just use the roads and the footpaths."*

Two people who identified rivers as edges also defined them as pathways.

Edges: A main characteristic of edges are that they (generally) cannot be physically crossed, but in the environment used any feature could be passed through as collision detection was not implemented. Therefore, the test subjects were asked if they noticed anything that they felt they should not be able to cross as opposed to explicitly a feature that restricted their movements within the environment. The main feature identified in this case was the steep slopes feature (37% of people) which is a grouping of all features that would impose movements owing to a steep incline or decline such as cliffs, mountains and sharp drops. This is closely followed by the river features with only one less person identifying it as an edge giving it a total percentage of 33%.

Districts: In an urban environment, districts are generally defined by a key characteristic or activity. Obviously in a rural environment with far fewer features the items that can be used to define a district in turn become fewer. What is seen from the interview results however is that the two top districts mentioned have a unique visual in the natural terrain and have a specific use. The urbanisations (33%) stand out owing to the manmade appearance, and the plantations (30%) were linear planting of trees which made them stand out against the random placing of the natural vegetation.

Nodes: Nodes were described as key decision making points either on the user's route or any other route through the environment. The first item, river crossings, does not refer to the crossing of two rivers but the user's movement crossing the river. The main aspect about the crossings identified was with respect to how many times it was crossed and in what time frame with responses such as:

> *"the number of times I crossed over things like the river"*

> *"the river where I crossed it twice in quick succession"*

"where you cross the river twice in a loop"

Road junctions on the other hand were where two roads crossed, although mentions of intersections between explicit road types were not aggregated into the road junction classification. This was owing to that in the two interviews where these were defined this way one explicitly only described junctions with main roads until river and road intersections were mentioned when general roads were used, whereas the second spoke of an explicit junction:

"crossroads between the main roads, and the main and minor roads, and then the crossroads between roads and rivers."

"on the Road to Hueli and there was the pathway crossing the road"

The mention of roads was of no specific aspect about them, but the mention of road bends explicitly referred to particular points along any road that could be corresponded to the map.

The features not making the top four of the Node category make up some 60% of all feature mentions. This, along with the distribution descriptive indicates that there were several other features mentioned with relatively large occurrence counts.

Landmarks: Interestingly, although landmarks are generally point features in this study it has been found that the top two features identified as landmarks are the linear features of roads (33%) and rivers (30%). In one interview it was mentioned that it was the location of the feature that identified whether the person was in the correct location, or the general direction and shape of the feature corresponding to the map:

"when I was going over certain places that things should be on the right hand side or left hand side"

"If it was turning towards the north you could see compared to the map were you were, and then looking around, if you could see where it had come from"

The points of interest associated to the info-marks were also mentioned as landmarks by seven people (23%) with an additional person using a point-of-interest marker on the map even though they did not have the info-marks displayed to them, stating:

"The only thing I did look at was Shale, and I thought when I came across a load of white things I thought 'I wonder if that's what shale is'"

Improvements: Within the interviews, users were asked what could be improved with the environment to make the task easier. 24 out of the 30 test subjects were asked to identify anything they would like adding, changing or removing from the environment, with nine people stating that they would like textual information in the environment identifying specific features. Seven identified that providing road signs placed at road junctions and path turning points showing the direction to nearby settlements or points of interest would be beneficial, and four wanted some form of feedback about how far they had travelled or at what speed they were moving at. Other responses included having physical features at the location of the points of interested associated to the info-marks and at route turning points, prompts indicating proximity to features (either textual or real world sounds originating from the feature), the display or emphasis of additional features on the map, and displaying additional information about current altitude. Table 31 highlights the improvements mentioned by users with the four distinctive types identified.

Trial	Improvements	Type
1	Buildings in the environment	
4	Names of places	Location
	More obvious rivers	
	Orientate head with ground surface	
7	Urbanisation names above places	Location
10	Distance display	Distance
	Physical feature at info-mark points	Location
	Place names - Hotel Rosa	Location
11	Show fences on the map	
	Buildings on map	
12	Alter the route	
	Landmarks at turning points - towers	Landmark at turn
13	Place names in the environment	Location
	Building and farm names in environment	Location
	Welcome to signs	Location
14	Map change	
	Direction pointers to places - floating	Location
15	Signposts at turning points	Direction
16	Labels on features	Location
17	Distance travelled	Distance

	Road signs at junctions	Direction
18	Different vegetation depending on elevation etc. - more realistic	
19	Speed display	Distance
	Directional signposts to settlements	Direction
	Signs indicating place name	Location
20	Altitude	
	Scale - travel speed	Distance
	Sign posts on junctions	Direction
21	Sign posts on junctions	Direction
	Sign posts at route turning points	Direction
	Sign posts at route crossing roads/rivers	Direction
22	Road intersections at route turning points	Landmark at turn
23	None	
24	Indicator in the environment showing POI location	Location
	Signposts indicating direction to POI	Direction
	Prompt indicating road/river proximity	
25	Signs indicating place name	Location
26	Audio cues - river sounds etc.	
27	None	
28	Farm names on the map	
29	Road signs indicating direction and distance at junctions and regular intervals	Direction
30	Legend on screen identifying road types	

Table 31: Improvements to the environment

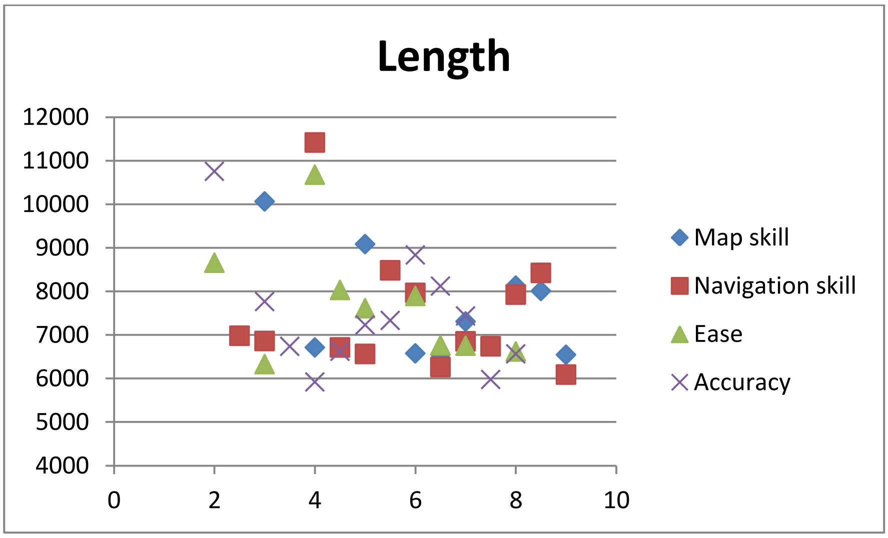
6.3.3 Comparison of Quantitative and Qualitative Values

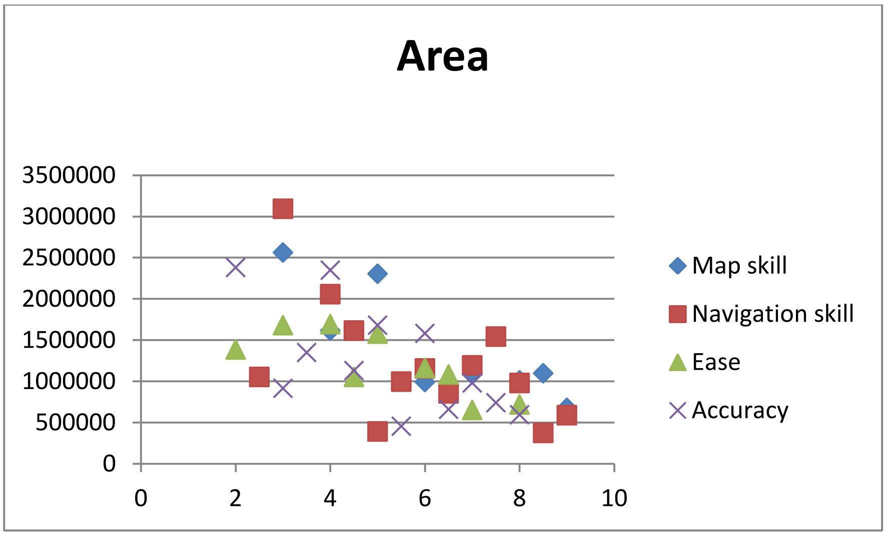
Additional analysis was performed via a comparison of the quantitative track log data with the qualitative ranking values obtained from the post-trial interviews (see Appendix 2 for the full table of data). Ranked Spearman tests were conducted for each qualitative/quantitative metric pair and significant results (using two-tailed testing at the 5% confidence level) were identified in the area metric for map skill ($p=0.007$), navigation skill ($p=0.006$) and ease of task ($p=0.014$). Significant results were also found in the combined metric in the ease of task ($p=0.011$) and accuracy qualitative metrics ($p=0.014$) as can be seen in Table 32. For all of the significant results a negative correlation was identified (a negative rho value) which signifies that as the rank for the qualitative metric increases the quantitative metric value decreases. This means that when users felt they had greater skill or that their performance was better, this was reflected in a better performance in the task with respect to the area and combined metrics in the circumstances identified.

	Skill with maps		**Navigation skill**		**Ease of task**		**Accuracy**	
	rho	p	rho	p	rho	p	rho	P
Length	0.017	0.930	0.053	0.780	-0.402	0.030	-0.163	0.389
Area	-0.483	0.007	-0.489	0.006	-0.320	0.090	-0.444	0.014
Time	0.201	0.287	-0.005	0.977	-0.414	0.026	-0.099	0.604
Speed	-0.260	0.165	0.058	0.759	0.299	0.116	0.025	0.895
Dist. from end	-0.344	0.063	-0.124	0.515	0.194	0.314	-0.110	0.563
Combined	-0.353	0.055	-0.290	0.121	-0.463	0.011	-0.443	0.014

Table 32: Spearman rank test results of qualitative and quantitative metrics

Graphical representations of the data in the form of scatter plots can be seen in the following six graphs (Figure 40).





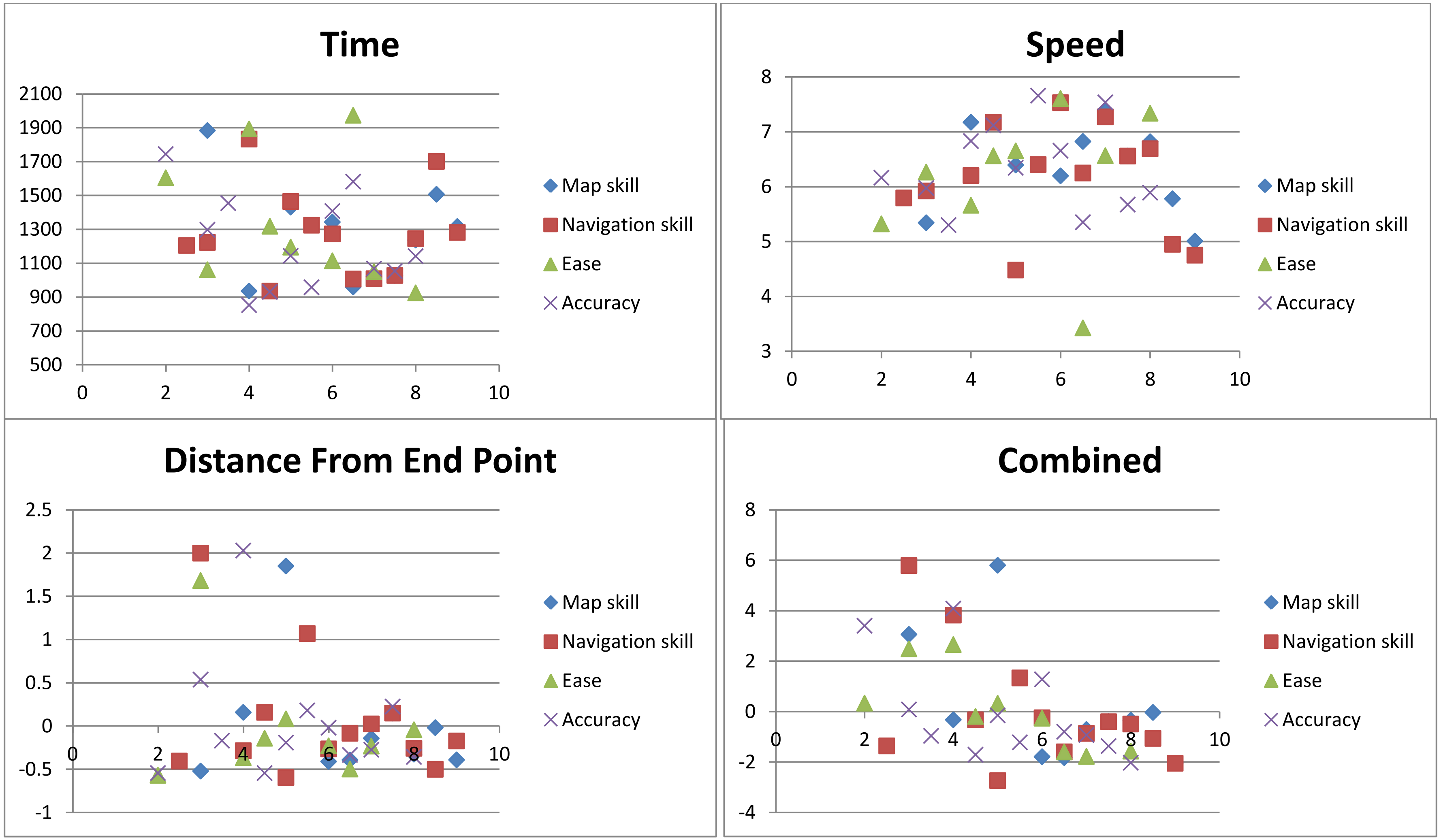


Figure 40: Qualitative against quantitative data analysis

6.4 Findings

By analysing the qualitative and quantitative data collected the research questions posed regarding the navigational and Urban Image aspect of the research can be answered.

6.4.1 Does Providing Geo-Referenced Contextual Information Aid in Navigational Tasks Within a Virtual Environment?

To answer the research question *“Does providing geo-referenced contextual information aid in navigational tasks within a Virtual Environment?”* info-marks were presented to the user of a VE in the form of text based information that appeared either onscreen or via a PDA device when the user was within one kilometre of the feature itself. The user was asked to trace a route within the environment as accurately as they could (the route being shown on a paper map) and track logs recorded of each user’s movements. Finally questions were asked in the post-trial interview regarding how easy they perceived the task to be and how accurate they believed they were. Questions were also asked regarding what features were used in the environment, whether the presentation of the text based information helped, and any improvements that could be made to the environment to make the task easier.

To identify whether the presentation of the info-marks had any recordable effect in the route tracing task several metrics were analysed from the track logs. It was found that the only aspect that appeared to be influenced by the inclusion of the info-marks was the average travel speed of the user, indicating that when a user was presented with the info-marks via a PDA they travelled faster through the environment. As no significance was found in duration or travel distance, this result seems to indicate that when the info-marks were presented via the PDA the users did not pause as many times or for as long as when they were not present, or interestingly when the same info-marks were presented onscreen. The difference between PDA and onscreen delivery methods could be possibly accounted for by the clarity of the text (as it was more legible on the PDA) or the movement of attention between the VE and the real world testing room, although the definitive reason is unclear. Also the use of an audio alert through the PDA delivery method may have had an effect on users noticing a change in the info-mark being displayed. It should be noted however that although statistically there was no evidence of any other effects of providing the info-marks, there could be a number of type 2 statistical errors (false-negatives) owing to the relatively small number of observations in each treatment group. This is a critical issue with the study presented as with such a small sample any statistical evidence that does appear would be questionable due to such statistical errors. In the context of this study, it was not possible to obtain a larger sample size in the time available and so any statistical findings should be used with caution. However, the primary aim of this study was to develop and implement the technology so that future studies would have a starting point for such investigations. Additionally within several of the groups outliers were identified within the boxplots which may have an influence on the overall significance between tests.

Although little evidence of the benefits or detriments of providing info-marks were identified from the quantitative analysis of the tracks, qualitative results from the interviews indicate that when the info-marks were presented to the user via a PDA, they felt that the task was both easier and that they were more accurate. Again however when delivered by the onscreen delivery method there was no significant differences in both cases.

When comparing the qualitative ranking scores (self reporting of map and navigation skill, and the perception of ease and accuracy) were plotted against the track log metric data, it was found that there appeared to be a negative correlation between the area metric and the ranks provided. This can be interpreted as indicating that when a person perceives their skill and performance to be high, the area between the intended and actual route is smaller.

With regards to the use of the info-marks, all but one user who had them present indicated that they made use of them, and the one who didn't identified that they were useful but they did not rely on them. When asked how they were used the responses indicated that they were used in much the same way as landmarks – identifying the user's current location, confirming whether they are on the correct route, or a combination of both. One user when asked whether they made use of the text that popped up on the PDA stated *"Yes I did, in fact they are like landmarks to me."* Additionally, feedback from the think-out-loud analysis identified the use of the info-marks as a means of confirmation of location. When the info-mark display changed, responses such as *"that's a good sign"* were given. When looking at the in-trial responses it is also possible to see that the info-marks were mentioned more when presented via the PDA delivery method. This could be because of the noise made when a new info-mark was displayed prompting a response from the user rather than simply noticing the change in the on-screen trials. From the in-trial feedback it is apparent that although roads are used heavily in the navigation process, when the info-marks are displayed the use of rivers decreases. This could simply be owing to the intended route not travelling in close proximity to rivers as it does with some roads and so the info-marks become more useful to know when to turn replacing the need to confirm travelling in the correct direction through observing the river.

With relation to the sub-question "*Does the method of presenting info-marks make a difference to their effectiveness?"*, results from the qualitative and quantitative analysis have shown a difference between how people felt they performed in the tasks and how they actually performed. The data have shown that there was a difference between delivery methods with regard to average travel speed, and that when delivered via the PDA users felt that the task was easier and they were more accurate.

6.4.2 What Features in a Rural Virtual Environment can Take the Roles of the Urban Image Theory's Five Feature Types?

To help identify what features take on the roles of the five UIT feature types (path, edge, district, node and landmark) users were asked to identify the features in the environment that fit with some of the main characteristics of the specific type. From the results it was clear that there were features that could take on type classifiers, but in many cases a feature was identified in multiple classifications. For instance, the rivers within the environment were highlighted as being both paths and edges, although the edge aspect was predominantly owing to how they would be interacted with in the real world.

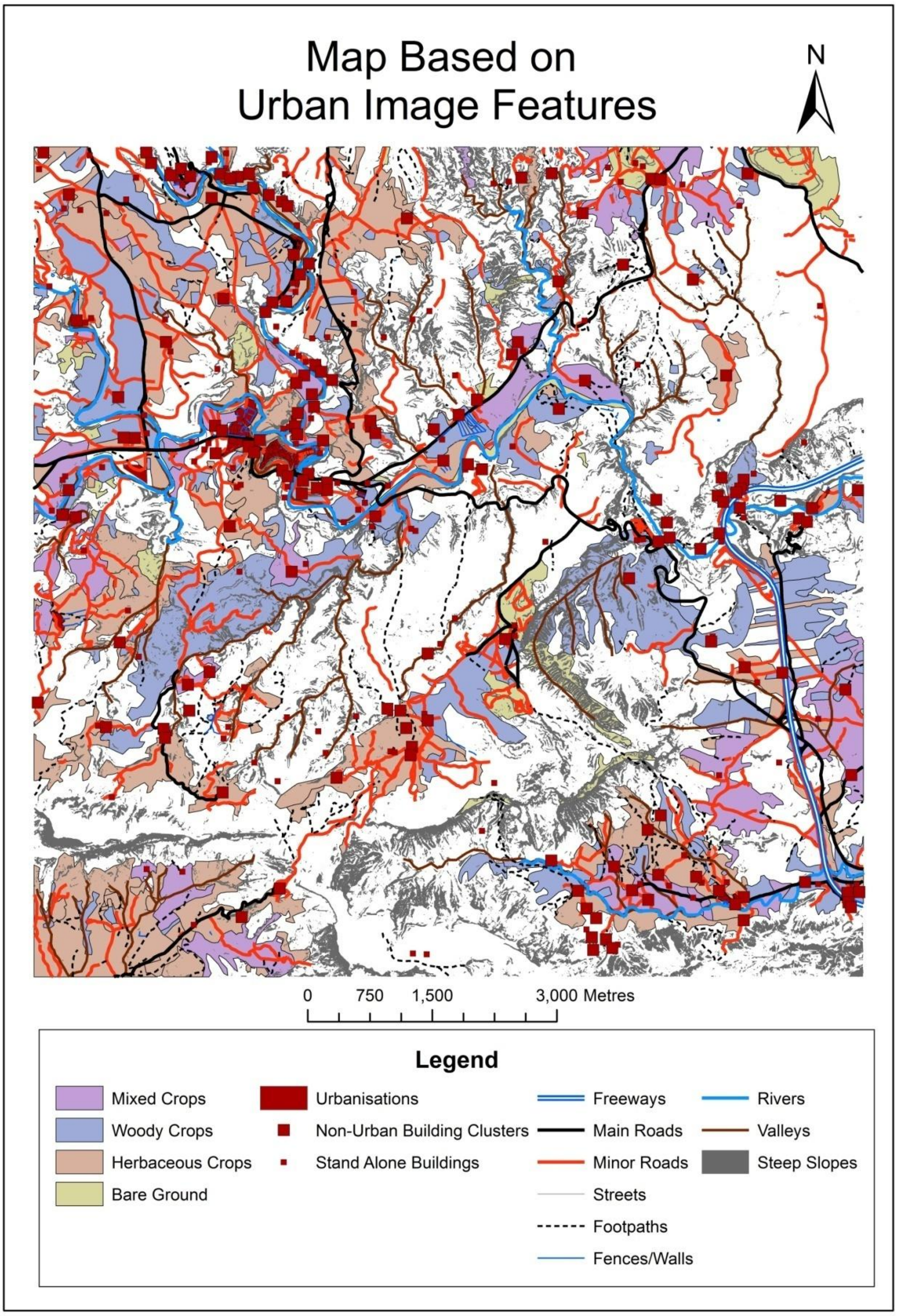


Figure 41: Map of UIT features

By identifying the features that people felt were in each classification a map can be devised that may be useful for navigating in the environment by highlighting features that people noticed (Figure 41). Also by looking at the map specific areas can be seen where classified features are scarcer and so if additional items were to be put into the environment then the areas they were needed can be identified. For example, owing to the different vegetation types being seen as districts, valleys being pathways and steep slopes being edges, these features can be highlighted on the map so that if a person is in a location where manmade structures are not visible these natural aspects can be used. On the map produced a diagonal strip can be seen running from the top right corner down towards the bottom left where smaller numbers of features that can be classified appear. Therefore along this section it may be of benefit to include other features that could be used to aid in navigation. These features would largely depend on the context of what the environment was being used for, but could for example be artificial structures placed such as buildings or roads.

It is worth noting however that the context that these classifications are being used in this study is that of landmarks in VEs discussed by Vinson (1999) and as such may not be the same as the UIT inside urban areas. In an urban context, a person should be able to identify which district that they are within, whereas in the VE here although plantations were described as districts it would most likely be difficult to identify which plantation you were actually within. However, in an actual rural environment in the real world this difficulty could be reduced as in the virtual representation the contents of the plantations (i.e. trees) have little variation between plantations – in the real world each plantation could have a different crop and a different structure.

6.4.3 Other Findings

As well as the primary research questions, other findings have been identified from the data analysis. Participants were asked what changes they thought could be made to make the task easier. Responses given could be split into four main categories: providing information regarding the direction of features or locations, displaying the names of places or placing a physical object to identify the physical location of info-marks, providing feedback about the travel distance or speed, and placing a physical feature at the route turning points. The latter of these is based on this specific task and as such would not be of benefit in other environments, as well as the features at info-mark locations. However, the other items may be of benefit to other VEs as they revolve around aspects that could be present in any environment and not specific to this task. Therefore, it may be of benefit to include features such as directional signposts at road intersections, the names of villages or buildings either on the structure or on entrance roads, and some feedback as to how far the user has travelled or a speed indicator.

In addition to this the study presented here has seen how the inclusion of the info-marks has an effect on which other features in the environment are used to navigate, with the inclusion of info-marks appearing to decrease the reliance on rivers to identify location and destination. This means that the inclusion of the info-marks could be of particular benefit in environments where such features are not present.

6.5 Summary and Conclusions

Results collected from both qualitative and quantitative sources have provided a mixed picture of the effectiveness of providing the info-marks to users on their route tracing performance. Although statistically little evidence of effect has been identified from the track logs collected (with only travel speed being shown as decreasing when info-marks were presented via the PDA), feedback from interviews identified that people felt more confident with their performance when the info-marks were presented. Additionally, the in-trial-feedback responses appear to show that when the info-marks are presented to the users they rely less on the river features found in the environment. In all cases roads were heavily used to help identify current position and target location. A particular interesting aspect identified was that if there was any significance identified either from the track logs or interview/in-trial feedback responses these only occurred in the trials where the info-marks were delivered via the PDA. When the info-marks were delivered via the onscreen method it was found that the users did not statistically feel any more confident in their accuracy or the apparent ease of the task. Therefore there appears that the delivery method may have an influence on how effective the info-marks are.

Within the interviews, data were collected as to which features in the environment could fit into the different UIT classifications as Vinson (1999) stated that features falling into each of these categories should be included in VEs. Feedback highlighted several common feature-classification combinations such as roads being pathways and boundary features (fences and walls) being edges. There were some more interesting combinations identified however such as vegetation types being highly rated as district features, along with valleys featuring as pathways. Other features were also identified which would not generally occur in urban regions but may be more frequent in rural areas such as steep hill slopes and rivers. Nodes were generally intersection points not only between roads but also between rivers and road/river intersections. These different features highlight that when a VE is created of a rural area, the same UIT classifications can be used but with different features than those that are more at hand in an urban setting.

With regards to improvements that could be made, two predominant aspects identified in the interviews conducted were the presentation of information regarding travel speed or distance travelled, and the inclusion of signposts within the environment. Therefore it can be concluded that in such rural VEs it is important to include these visualisations to aid the users in the navigation within the environment.

The main contributions from this study were the methods implemented for both the display of the info-mark data and the collection of the track log information. By using current mobile technology in the form of imitating a Bluetooth signal transmitted from a GPS receiver it was possible to integrate field technology (in the form of a PDA) with a virtual representation of a real world place. By doing this the same delivery method can be used in the field meaning that a person could be trained in the VE to use the equipment before its use in a real world application. Although track logs have been used in previous studies (i.e. Chittaro et al. (2006), Drachen & Canossa (2009a), Hoobler et al. (2004)) the use of both quantitative results and qualitative feedback is a new implementation. It is important to perform both forms of analysis as although effects may not be present in one set of data they may become apparent in another (as seen in this study).

In final summary, this study has identified that although statistically there is little evidence of the presentation of info-marks having an actual effect on user performance in the route tracing exercise (and when they did it was only when delivered through the PDA) it appears that their presentation does make users feel more confident in their performance and that the task was easier. However, this appears only to be the case when they are presented via a PDA so the delivery method could have a major effect on their usefulness. Also, some correlation was found between the area metric used for analysis and the self reported map and navigation skill / perceived ease and accuracy of the task. This appears to indicate that when users felt more confident in their skills and performance, the area metric decreased. This was the only metric however that showed such a relationship with the self reported ratings. Additionally it has been shown that there are several features that can fit into the UIT classifications and so can be used by a developer to enhance navigation performance. Finally interview responses have indicated that users feel that feedback regarding to travel speed or distance would be highly beneficial as well as road signs within the environment.

6.6 Further Study

Although this investigation has provided evidence on the effects of info-marks and on the classification of different features into the UIT classifications, it has also highlighted that further study is warranted into several aspects to gain a better understanding.

6.6.1 Different Delivery Methods of Info-Marks

It has been shown that there appears to be a difference in effectiveness of the info-marks depending on their delivery method with delivery via a PDA emerging as more beneficial. Therefore further study is warranted into what makes this delivery method appear to be more useful and whether other delivery methods could be even more beneficial.

6.6.2 Interaction Methods for Feature Types of UIT

In this study, different features that can take the role of the UIT classifications were identified. However, in the case of the environment used, users were not limited to features they could cross (i.e. users were able to walk over rivers without penalty). In other implementations, if such restrictions were enacted, the identification of features such as edges may change. Additionally the method of locomotion could alter the identification of features. Here the user moved at a constant speed using a mouse interface without an avatar. If a similar implementation was used but a driving interface implemented (a steering wheel control or a driving type Heads Up Display) then users may interact differently with the system and stay on road features.

Therefore, the implementation of the environment with regard to user interaction may have an effect on the features identified in the UIT classifications. A future study could implement different interaction techniques (driving, treadmill walking, flying) and identify if the feature types would alter as happens in real world instances, and so determine which features may be more important to highlight on a map as a means of improving navigational performance.

6.6.3 Implementation of Travel Distance

A feature mentioned in multiple interviews was that the implementation of a tool to help identify distance travelled would be very beneficial. Such tools could be a speed display or some method of identifying distances travelled either as a whole or between way points which could be manually set by the user. The use of such tools could be analysed in a similar context of this study to identify if the feedback on this information does in fact have any beneficial effects on the accuracy of following a defined route, as well as identifying if a particular method of presenting the information is preferential to others.

6.6.4 Inclusion of Signposts

The final change identified from interviews was the inclusion of signposts within the environment, both at road junctions and where roads enter villages. Both of these are homogeneous to the road signs found whilst travelling along roads in real world environments which are used to identify the direction to travel in and to confirm that the navigator is in the correct location.

7 Part III: Visualisation of Track Data

Can a multi-metric analysis be performed on track log data to derive information about trial performance, how features within the environment are used in navigation tasks, and what features are present in the environment itself?

7.1 Introduction

In the previous study data were collected from users in the form of track logs as a means of assessing the effects of presenting geo-located information to users on their navigational performance with regards to a route tracing exercise. Analysis was performed using statistical analysis of metrics obtained from the track logs (duration, travel distance, average travel speed etc.) and qualitative analysis of interview data. Although these methods provide a suitable analysis framework for that investigation, the data collected can be further explored to gain additional understanding of user movements via visual analysis.

Many previous studies have used the visualisation of movement data (either in real time or as a post analysis tool) as a means of determining characteristics about the environment or user movements (i.e Chittaro & Venkataraman (2006), Drachen & Canossa (2009a), Hoobler et al. (2004)). Generally however, such visualisations only show single pieces of data (i.e. visited locations) at any one time or comprise of a relatively small number of movement variables (low number of users). This study however uses methods to display data from multiple datasets with regard to a larger number of observations (track logs), and so assesses whether showing multiple pieces of information at the same time can be of benefit.

In this chapter, methods for the extraction and visualisation of data from the track logs are discussed. Firstly, individual datasets that can be used as a means of assessing movements are gleaned from current track logs using GIS software. Once obtained, this data is presented in such a way that all datasets can be analysed simultaneously to gain a more insightful understanding of areas of interest within the datasets, before the resultant visualisations are compared with the VE itself to identify relationships. Finally the same data extraction methods are implemented on the separate groups of track logs identified by the type of info-mark delivery and then the datasets compared using the visualisation technique.

7.2 Data Extraction

From the track logs created we can extract several different informative datasets (track count, clustering, and congestion) as used in Chittaro et al. (2006). Manipulation was all undertaken in ArcGIS using a variety of pre-defined tools as well as custom coded methods implemented in VBA.

7.2.1 Track Logs

Initial display of the track logs can be generated by simply drawing the tracks, with colour coding implemented to differentiate between the info-mark display types if needed (Figure 42). A problem with this method is that it is difficult to determine how many tracks are in a particular location. For example, after the second turning a number of tracks converge but from the simple track display the number is impossible to determine without zooming in on the data. Additionally it cannot be determined whether people were travelling slowly or at speed, or what portion along their track they are at a specific location. This information can be obtained using additional data extraction and visualisation techniques as will now be discussed.

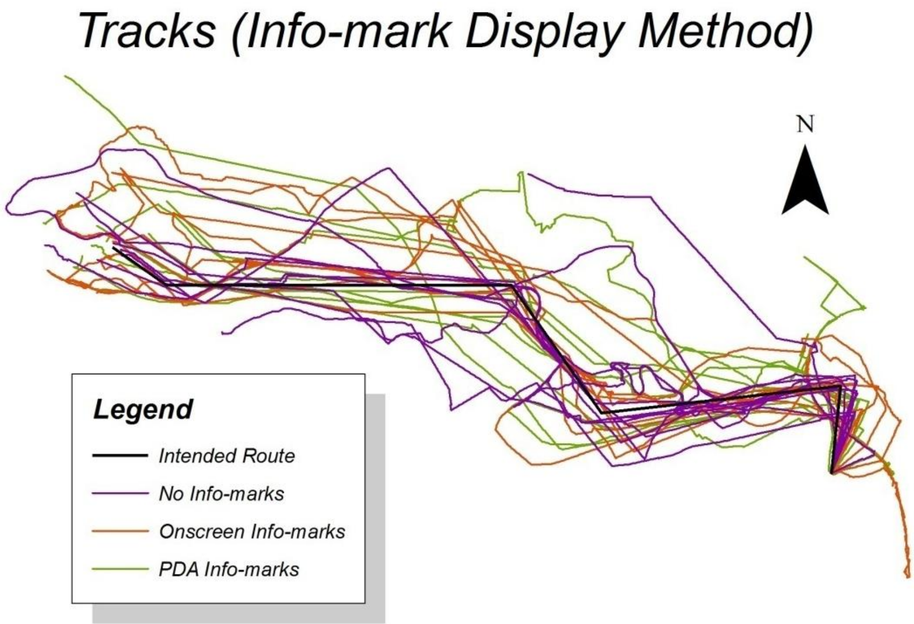


Figure 42: Track logs

7.2.2 Track Count

From the linear track display it is difficult to derive the actual number of tracks which are present in the vicinity of a specific location. This type of information is of interest as by identifying areas visited by multiple users, we can distinguish locations that for some reason draw users to them. These locations may be the intended route itself or key reference points that people used as a means of clarifying their location.

To analyse the data in this way, several of the ArcTool features were implemented in sequence using ArcMap's Model Builder. Firstly, each of the tracks was buffered at a specified distance (in this case 100 metres) using the buffer tool. This was required as using the linear paths themselves would identify very few places were more than two paths would intersect owing to the non-restricted nature of the environment. Once buffering was complete, the buffers were rasterised and then reclassified. The reclassification process ensured that the rasters would contain values of 1 where the buffer was present and 0 where it was not. By default, the raster representation was 0 for buffer present and NoData for where it was absent which causes problems in the subsequent stage where the raster cell values were added together. Finally, all of the raster representations were added together providing a single raster with cell values ranging from 0 to 30 which represent the number of tracks that crossed or were within the vicinity of that cell (Figure 43).

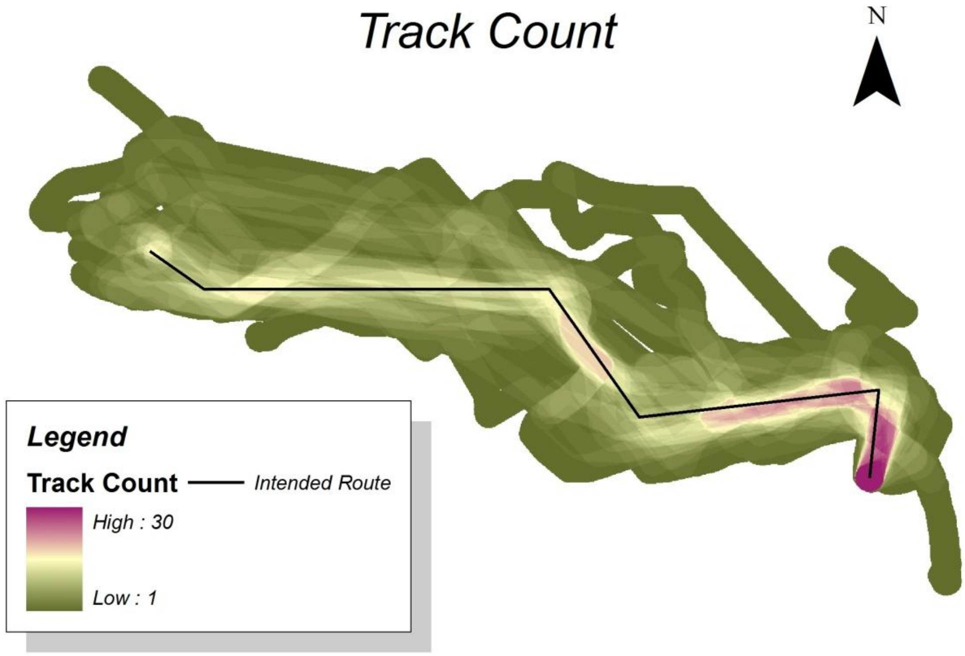


Figure 43: Track counts

7.2.3 Clustering – Waiting Points

Although the track count displays information about where users travelled to, it does not provide any information about their travel velocities and places where users paused to re-orientate themselves. Data extraction and analysis was performed to identify areas where users slowed their movements or came to a halt. These changes in velocity could indicate either a decision making point or an area where the user felt disorientated.

The clustering analysis performed used methods for programmatically comparing the distance between points in each of the track log datasets. Programmatic loops were implemented to scroll through each point in the log and determine the distance from it to every other point. If this distance was less than the specified threshold (100m) then an extra field in the shapefile dataset named ‘Count’ was incremented by one.

When all of the individual point shapefiles were merged into a single image (Figure 44) we can identify the locations that had a proportionally higher density of points which would correspond to the areas that one (or more) users would have slowed or stopped. The resulting image shows the level of clustering for each individual track and not the clustering of all tracks merged into one dataset meaning that a value of 200 would indicate that there are 200 points within the same track log that are within the specified distance.

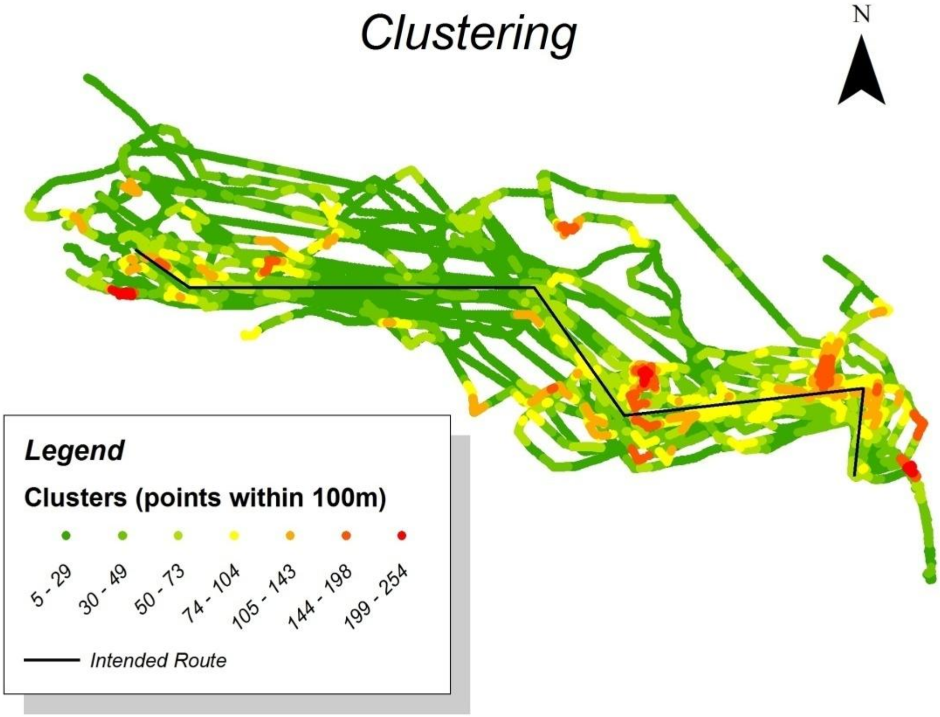


Figure 44: Clustering

7.2.4 Congestion

In this investigation, the term congestion indicates the relative density of points at specified percentages along the track duration. Congestion would usually infer that multiple people were in the same place at the same time, but owing to the nature of the experiment from which the data was collected users were only ever in the environment on their own and their total travel times varied. Therefore to assess congestion, relative times along tracks were used. This meant that rather than determining how many users were in a specific place 10 minutes into their trial, the number of people in a place at a time equal to 20% of their total trial time was calculated. The method implemented here assigned each point in a track log to a whole number percentage value ranging from 0 to 100. This meant that multiple points in the track would carry the same percentage value which also meant that there would be a match in all other track logs as well.

Once the points in a track had been given a percentage value, all point shapefiles were merged into a single large shapefile. Programming code was then implemented to go through the merged dataset and create new shapefiles containing all points of a specific percentage value, resulting in 101 shapefiles (0 to 100 percentages) being generated. Once these individual datasets were produced, a kernel density function was applied to each resulting in a series of diagrams indicating how densely clustered the points were at that percentage. All of these individual kernel density diagrams were compared and the largest cell value used in a final density diagram as can be seen in Figure 45. A drawback of this method is that when the final diagram is produced we get 'bubbling' within it which is caused by the discrete groups of points being used as opposed to singular items. In Figure 45 the data within 100m of the starting point were removed to ensure that the data in the rest of the visualisation could be analysed better as all users would have points of value 0 in the same location which results in this region having a considerably larger value than the rest of the data.

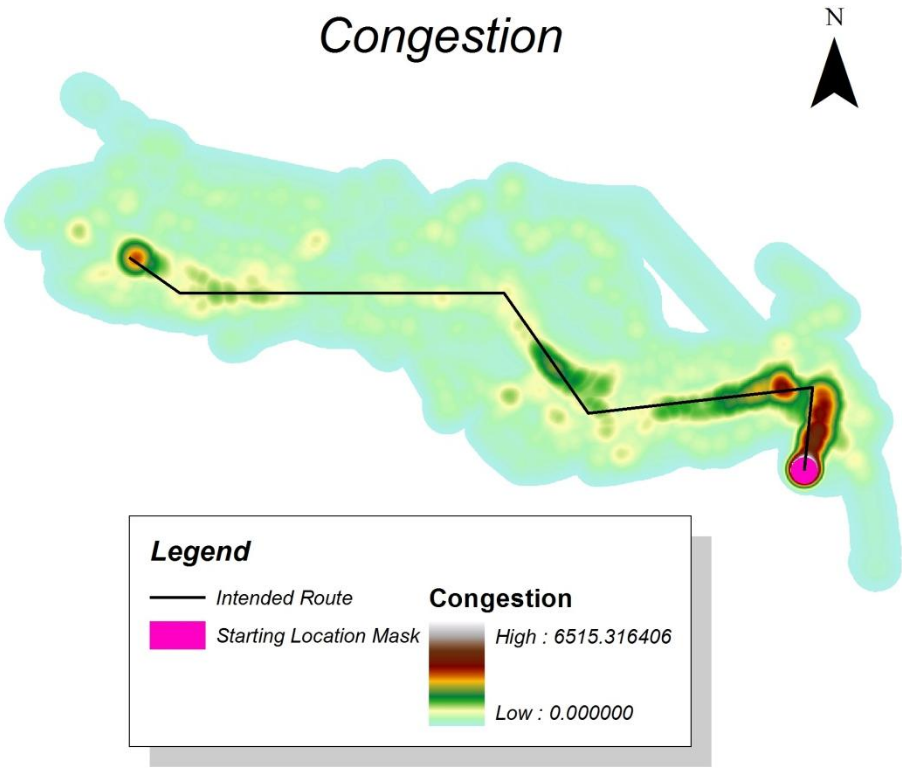


Figure 45: Congestion

7.3 Individual Analysis

One method of analysing each of the metrics individually is the heatmap visualisation as used in Drachen & Canossa (2009a) which was the method used in Figure 43 and Figure 45. In this method, the data is visualised with a raster representation containing cells of different colours and intensities which represent the values of the raster cells. In particular this method is useful for ordinal data where graduated colouring can be used to represent the data.

7.3.1 Track Count

By looking at the resultant image from the track count data extraction we can see that the predominant areas where multiple users visited was in fact along the intended route indicating an overall pattern of following the route as intended. If we were to look at all the linear tracks in one image it can be difficult to actually determine how many of the tracks are in any location, whereas by visualising the data in this way we can clearly see the more travelled locations. For example, we can see that in the central region there are a higher proportion of tracks within the same region having a track count value of 17-18 (Figure 43). Another feature that we see in the diagram is highlighted in more detail in Figure 46. Here we see a triangular shape appear above the intended route which implies that multiple tracks followed this same route for some reason. By comparing this to the features found within the environment we see that the left hand region of the shape follows a main road indicating that people would reach this road and then follow it. Similarly we see on the right hand side that the higher track count region is an extension of part of the intended route which ends at the river, again indicating that when users reached this feature they realised they had travelled too far and so changed their course. The increased level of track count can however have some attribution from where a number of tracks cross but not necessarily heading in the same direction. For example, although some users may follow the road, others will simply cross it and so at that intersection point the track count will be increased by one even though the tracks are not heading in the same direction.

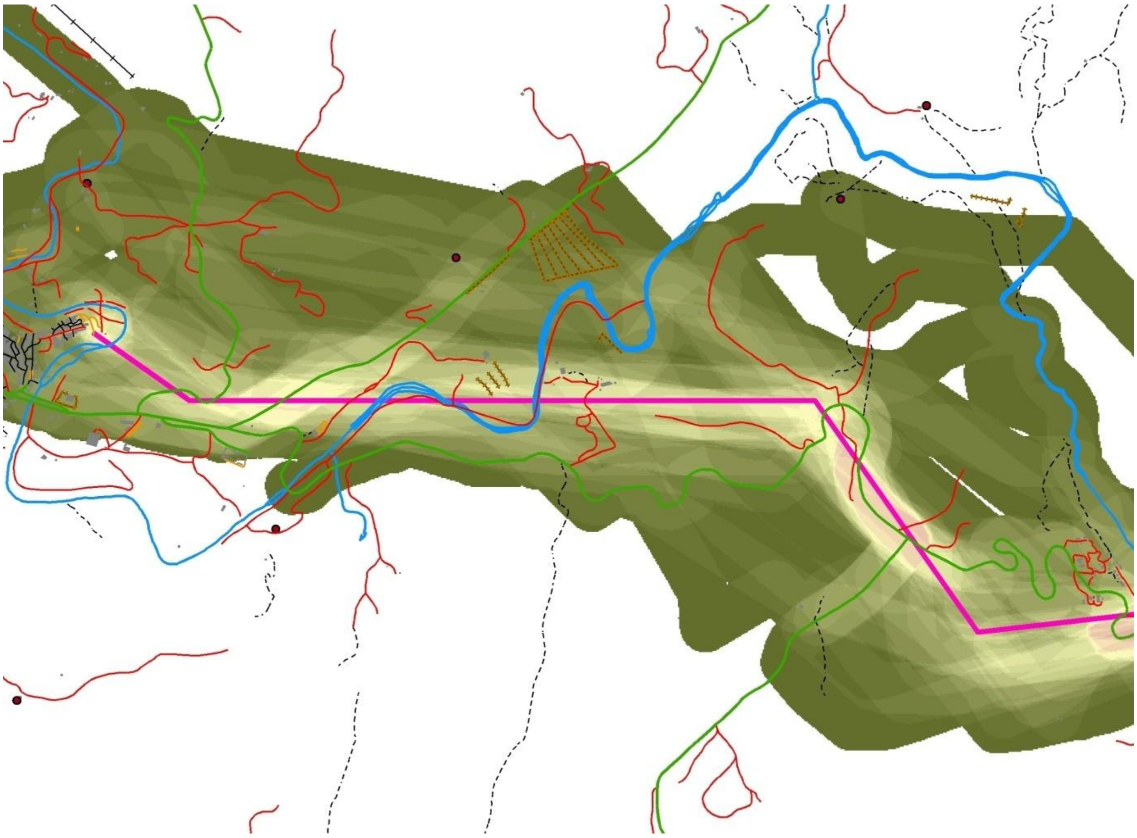


Figure 46: Track counts and environmental features (central region)

7.3.2 Clustering

From the clustering diagram produced (see Figure 47), several areas of higher level clustering can be identified (values of 144 – 254). The clustering locations indicate regions where the user slowed down or came to a complete halt. This slowing down may occur in the real world in high levels of congestion, if the traveller wanted to spend more time in an area (i.e. sightseeing), or if the traveller felt lost and so needed to re-orientate themselves. In the case of this experiment it is far more likely that the reason for slowing down and stopping is so that the user can look around to re-orientate within the environment. Therefore we would expect these regions to either be a place where the user finds it difficult to orientate or a place where the correctness of their motions has been placed into question and so they need to clarify their current location. An example of the latter scenario is if the user came across a feature which they were not expecting. These instances would indicate two different conditions within the environment: the difficulty in orientation would most probably be a result of a lack of distinguishable features; and the reassessment of location could be attributed to the occurrence of a distinctive feature that was not expected.

When compared with the environment we can see that several of the larger clustering locations (values of 144 – 254 points within 100m) match up with particular features (or areas with a lack of features) as can be seen in Figure 47. At locations A and D we see the clustering occurring where different features are within close proximity, with A having a single highway splitting into two to make a dual carriageway, as well as a minor road and some wire fencing. In D we observe the freeway with several minor roads and the river in close proximity. In C however we only see one feature – a winding main road. The reason for clustering here could be down to two main possibilities. Firstly, the user could be using the winding feature to determine their location which could result in a pause while the bend is identified on the map. Alternatively, the clustering could simply be owing to users following the winding feature itself meaning that more points would appear in the same location owing to the higher fractal dimension. Finally, location B highlights three places where waiting occurred where there were no features present. This means that the pauses are most likely occurring owing to the user having difficulty orientating themselves within the environment ready for a turn in the intended route.

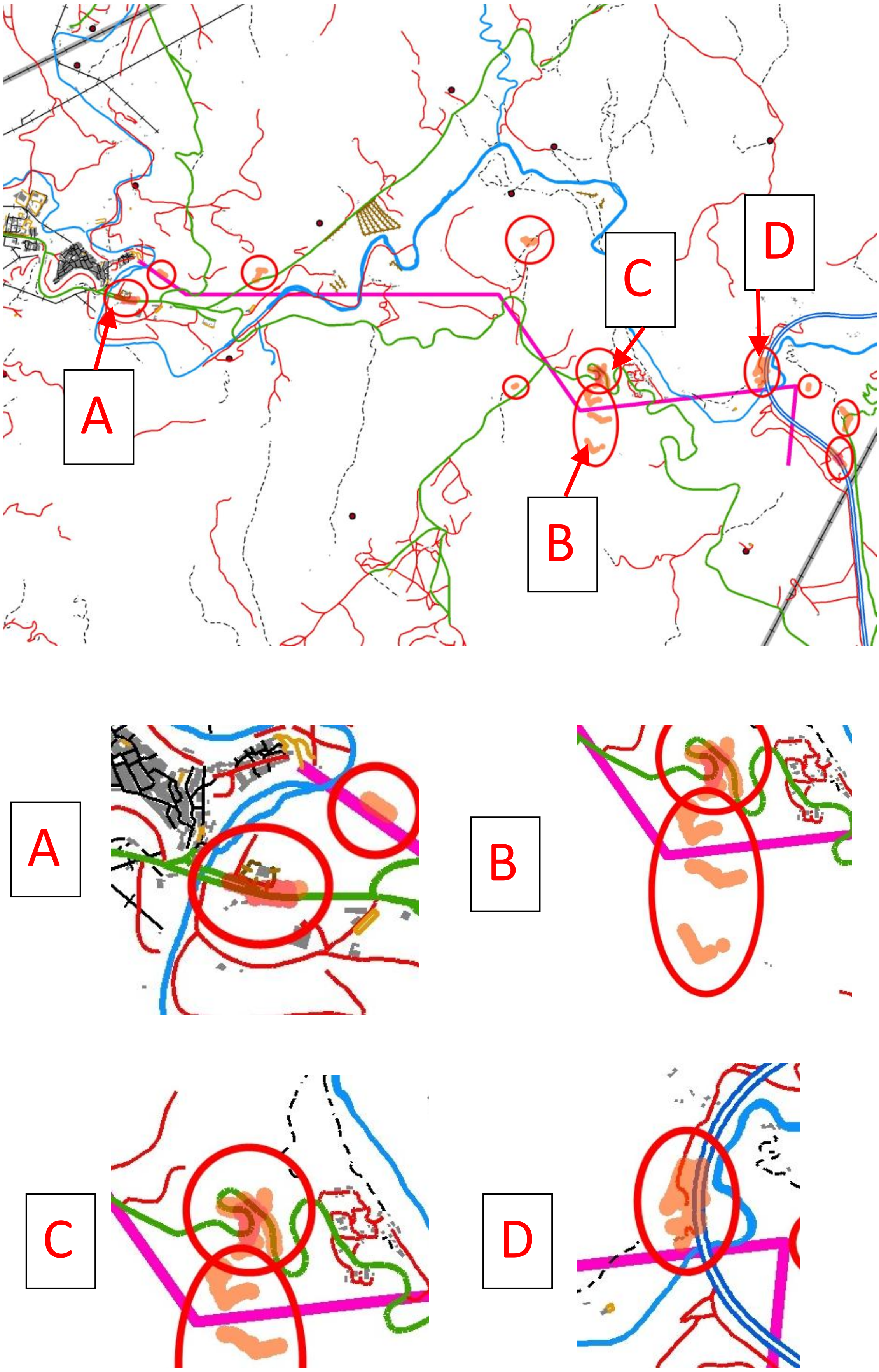


Figure 47: Regions of high clustering

7.3.3 Congestion

Owing to the starting location obviously being highly congested (everyone would have been in that vicinity at the start of their trial) when visualising the congestion in Figure 45 a buffer region of 100m radius from the start location was masked from the dataset. This allowed better display of the other areas with higher contrast than if the peak at the beginning was left in place. When any statistical analysis (or visualisations as discussed later) were used, the data from this area was not masked out.

Areas of primary congestion are near the starting location and the following region along the route, a central location, and at the end. Generally you would expect higher levels of congestion at the beginning as there has been less time for cumulative errors to surmount. In addition, you would expect a higher congestion near the end point as this is the place people would have been aiming to reach and as this would always be the last percentages of the overall track then users would roughly be within this region at the relative same time (unless they deviated from the route by a large amount and then ended in the wrong area). The main interesting location within the diagram is the central region where congestion increases. This increase indicates that there was some characteristic of the location in the environment that would draw a number of users to that place roughly half way through their journey. When compared with environmental features (Figure 48) we see that the highest amount of congestion in this central region falls at the intersection of two main roads suggesting that this feature is what was causing users to be in the same location at the same relative time.

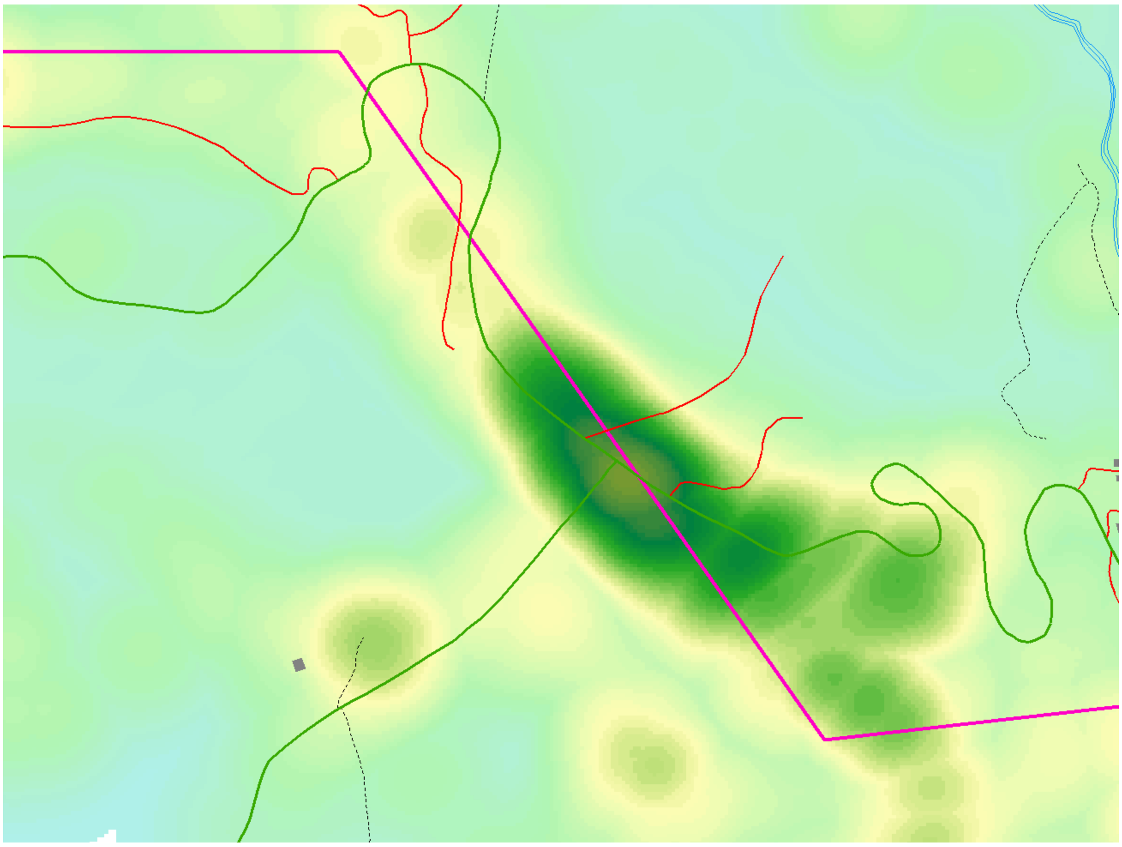


Figure 48: Central congestion region

7.4 Merging of Data

From the analysis of the previous visualisations we see some interesting areas such as the central region where congestion and track counts increases. However, it may be more beneficial to be able to determine whether there are relationships between the different visualisations and what these relationships may in turn signify. Although viewing the images generated side by side can give us some information about this, we are limited in terms of ensuring that we are in fact looking at the same location in all the images. This is not such a problem when you view the full extent of the images but as you begin to zoom in on areas it is easy to end up looking at different parts unless a method was implemented to ensure that you are always looking at the same location in all images. One possible method of comparing datasets would be to classify each of the calculated rasters into groups and then combining multiple rasters into one. However, visualising the information in this way produces an exponential number of final classifications used in the single raster image. For example, if we were to use the congestion and track count rasters we would first need to classify them into groups (30 classifications for the track count alone is simply far too many). If they were each classified into 5 groups, then the resultant raster would have 25 (5x5) different classifications. If another raster was added again with 5 classifications the resultant raster would have 125 (5x5x5) classifications. In this study both the congestion and track count were reclassified into five groups based on Natural Breaks (Jenks) which results in datasets described in Table 33 which can be seen in Figure 49.

	Values	**Classification**
Congestion	0 – 379	1
	379 – 984	2
	984 – 2317	3
	2317 – 5438	4
	5438 – 10163	5
Track Count	1 – 3	1
	3 – 8	2
	8 – 14	3
	14 – 22	4
	22 – 30	5

Table 33: Dataset classification boundaries

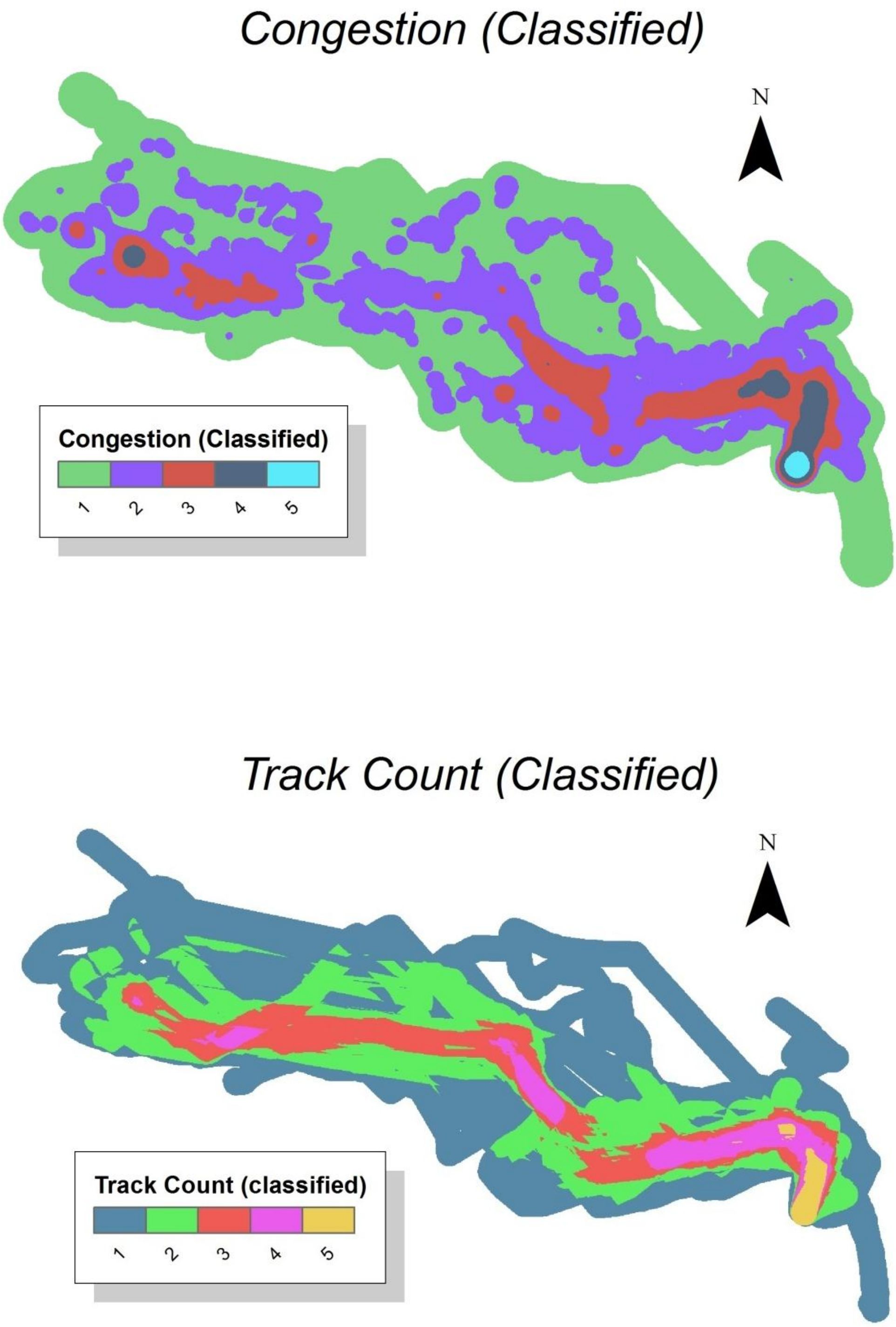


Figure 49: Classified congestion and track count datasets

Once classified, the two datasets were combined using the ArcMap raster calculator with the equation $[Value] = ([Congestion] \times 10) + [Track\_Count]$ resulting in a raster containing discrete two digit values whos first digit represented the congestion classification and the second digit the track count classification. After creation to allow for GIS visualisation the raster data was exported to a vector shapefile so that the combined value could be split into the separate components ([Value] / 10 for congestion and [Value] Mod 10 for track count) and then visualised.

Finally the different values are displayed using a different hue to identify the congestion classification and a different intensity to identify the track count (Figure 50). In Figure 50 if a cell has a classification of 4.1 (Light orange) this would signify a high congestion (4) and a low track count (1). In this example we can see that areas that had high congestion values also had a high track count (higher intensity of the yellow to red hues) as well as the lower congested areas also predominantly being in areas with low track counts.

Although this method can be used to display information from two raster datasets it would be difficult to successfully display more than this. If we were to add a third raster dataset, say clustering, we have few options on how to display this. One option would be to use patterns such as cross hatching or displaying all features as a regular grid of dots with spacing determined by one of the datasets (i.e. dots closer in more congested areas). This could however have a detrimental effect on how easy it is to differentiate between the other aspects used to show the other datasets. One solution is to make the image 3D meaning that the height aspect can now be used to identify a property.

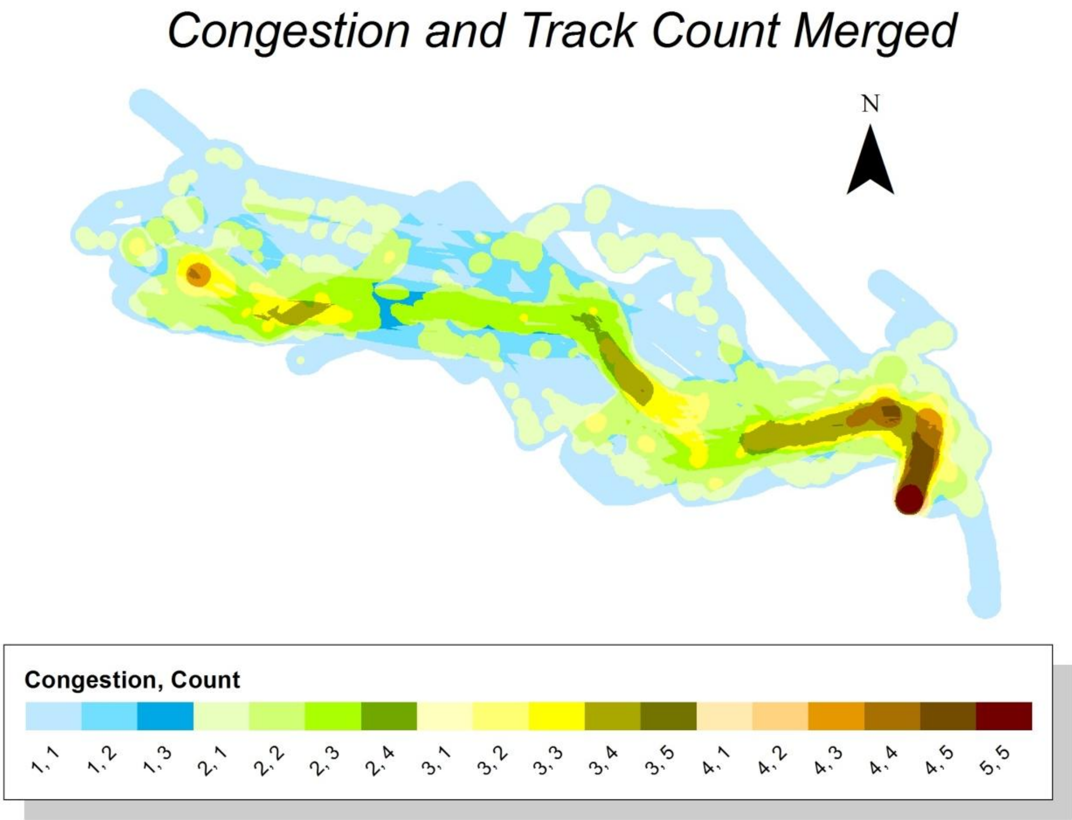


Figure 50: Merged dataset

To display the clustering information in the combined raster, an IDW interpolation based on the count field within the dataset was performed (Figure 51). To make the IDW display the locations correctly, points were added from the boundary of a 100 metre buffer around all points where the count value was set to 0. It should be noted that an IDW may result in the loss of some information from the clustering point data. For example, if there is a waiting point of one track surrounded by fast moving points from other tracks, the waiting point may be overlooked owing to the proximity of the large number of faster points. In this study however this does not appear to be an issue as in Figure 51 there are no high clustering areas which do not feature in both the vector point and raster IDW. As this data is being used for elevation, classifications are not needed as an explicit hue-intensity combination is not required, thus the continuous data can be used.

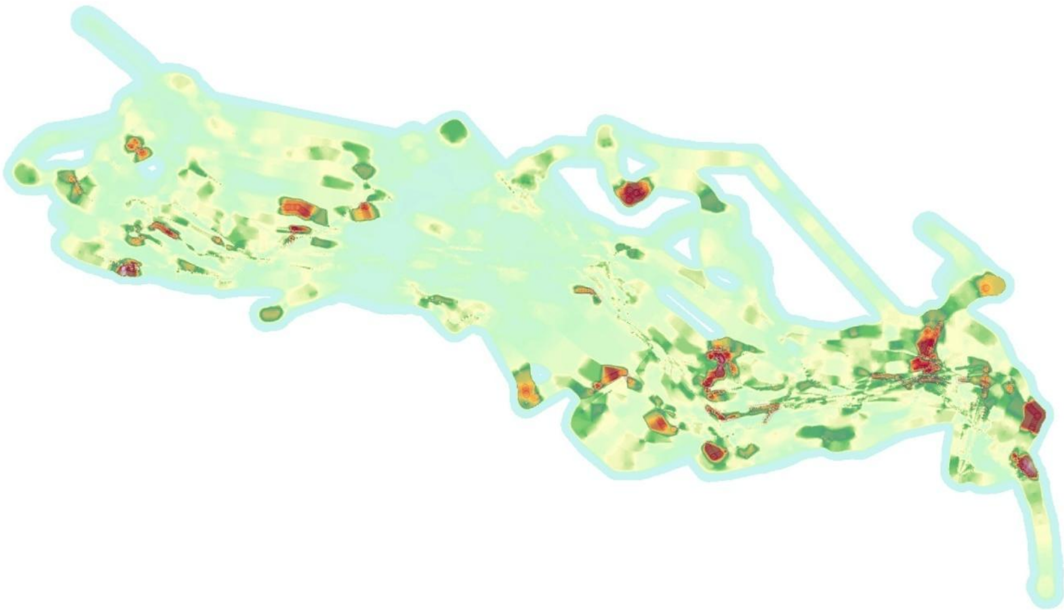


Figure 51: IDW diagram of clustering

7.5 3D Display of Merged Dataset

Once data had been extracted and merged as described, initial analysis was performed using ArcScene, a 3D GIS. By displaying the combined congestion and track count dataset with base heights derived from the clustering dataset we arrive at a visualisation as can be seen in Figure 52.

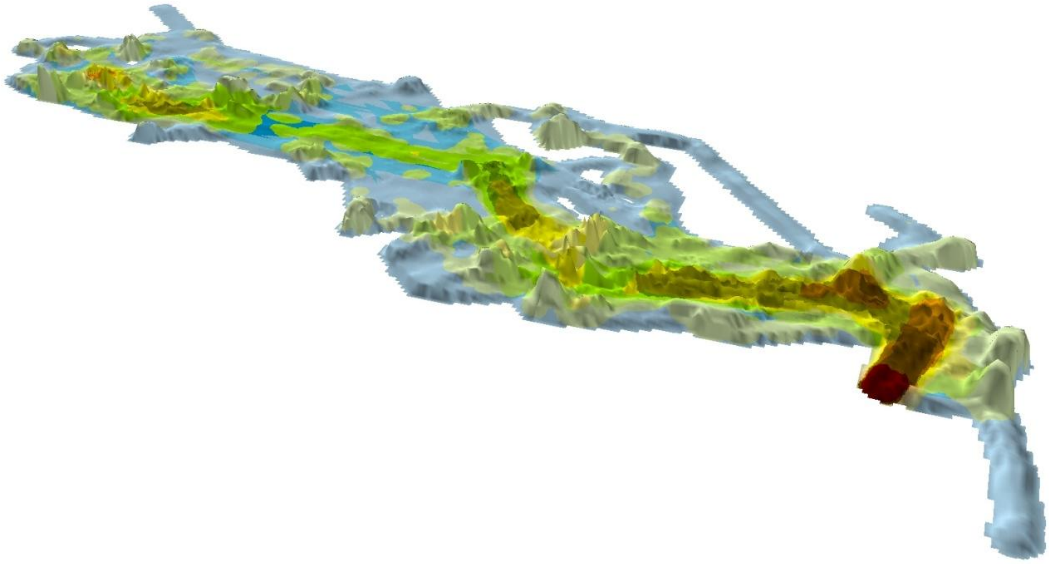


Figure 52: 3D visualisation of data

7.5.1 Special Features

From using clustering data as the elevation model, we find several feature types which can be used to describe particular aspects of the trial progress (see Table 34). Each of these types can be seen as a real-world topographical feature which in turn can aid in the description of the information to someone who is not trained with the visualisation used. For example, describing the beginning region to someone as a mountain range immediately invokes the idea of a region of peaks signifying higher levels of clustering throughout the region. As well as mountain ranges, individual peak features which are single peaks within a flatter data area can be found. These single peaks generally occur when a single user pauses for a long time in a region which other track logs have not travelled through. If a number of other users did travel in the same location, the peak would only form if a number of them paused in the same location as the IDW method would make the peak narrower if less clustered points surround it. Similar to the mountain range, a ridge is again a series of elevated peaks. Unlike the mountain range however the ridge tends to form suddenly and is a quite narrow linear feature. This can be caused by multiple tracks travelling in a wide area suddenly pausing along a linear feature in the environment, such as a road. The trench feature is closely related to the ridge feature and is signified by a steep valley running along the centre of the ridge formation. This would usually occur if the linear feature that caused users to pause in the formation of the ridge is then used as a pathway to travel along resulting in less clustered region owing to faster travel speeds. Plateaux features are generated from a wide area of constant velocity which produces a low level of troughs and peaks in the terrain. The valley feature is a linear region of low elevation between two higher elevation regions. These higher elevations could be formed from data features such as mountain ranges, ridges or even the edges of plateaux features, but faster movement through the environment forms the valley itself. There are two main ways in which valleys can form: firstly users travelling along a linear feature in the environment can generate them, and secondly they could be formed by a region where tracks are dispersed but still travelling at a higher velocity. These valleys in turn can be seen as parallel or perpendicular with parallel being generated in the former of the two formation methods where the valley follows the tracks and perpendicular being generated in the latter where the valley is across the tracks. For the remainder of this thesis, all topographic features relating to the visualisation of the track log data are prefixed with ‘data-‘.

Feature	Characteristics	Cause	Example (Figure 53)
Mountain Range	Collection of peaks in an area resulting in a generally elevated region	Group of waiting points or general slow movement	A
Peak	Single peak amongst a flat area	Single point of waiting	B
Ridge	Narrow line of peaks	Reaching of a linear feature that causes people to stop for re-orientation	C
Trench	Narrow valley through the middle of a ridge	Movement along the linear feature that caused the re-orientation which generated the ridge	D
Plateaux	Large area of flat terrain	General movement through the environment with a constant average speed	E
Valley	Area of lower elevation than surrounding regions	Faster movement with few or no stops between regions of slower movement	F

Table 34: Data features

The other aspects visualised (track count and congestion) are also important as they can be used to highlight any of these features which may be occurring in a more unusual case. We would usually expect data-peak features to appear in regions with low track counts as this would symbolise that possibly only one person found themselves needing to pause at that location. If we see the same feature appearing in a high track count region however this would signify that a large number of users were all pausing in almost the same location indicating some particular feature in the environment that caused the pause. Similarly we would also expect data-ridges to appear in lower track count regions as these generally occur when a person arrives at a linear feature which can be come across from multiple locations by different users. The relationship between congestion and track count is also important as high levels of congestion would not occur in regions with a low track count. A low congestion value and high track count indicates that users went to the same location but at different times during their trials so this can be used when making modifications to aid in the navigation process. For example we would generally want to have high track counts and high congestion throughout the route as this would signify that users all went in the same direction without deviations. Therefore, once changes had been made to the environment, track logs could be collected again and analysed based on congestion – if congestion values diminish along the track then people are deviating from the intended route. In the case of this experiment analysis was performed using the percentage of trial time rather than the absolute time. This meant that if a person travelled particularly slowly but along the same route as everyone else who all travelled at a uniform faster speed, the slower track would have the same impact on congestion as if it were travelling at the same speed as the others. The benefits of using this method were discussed in more detail in section 7.2.4.

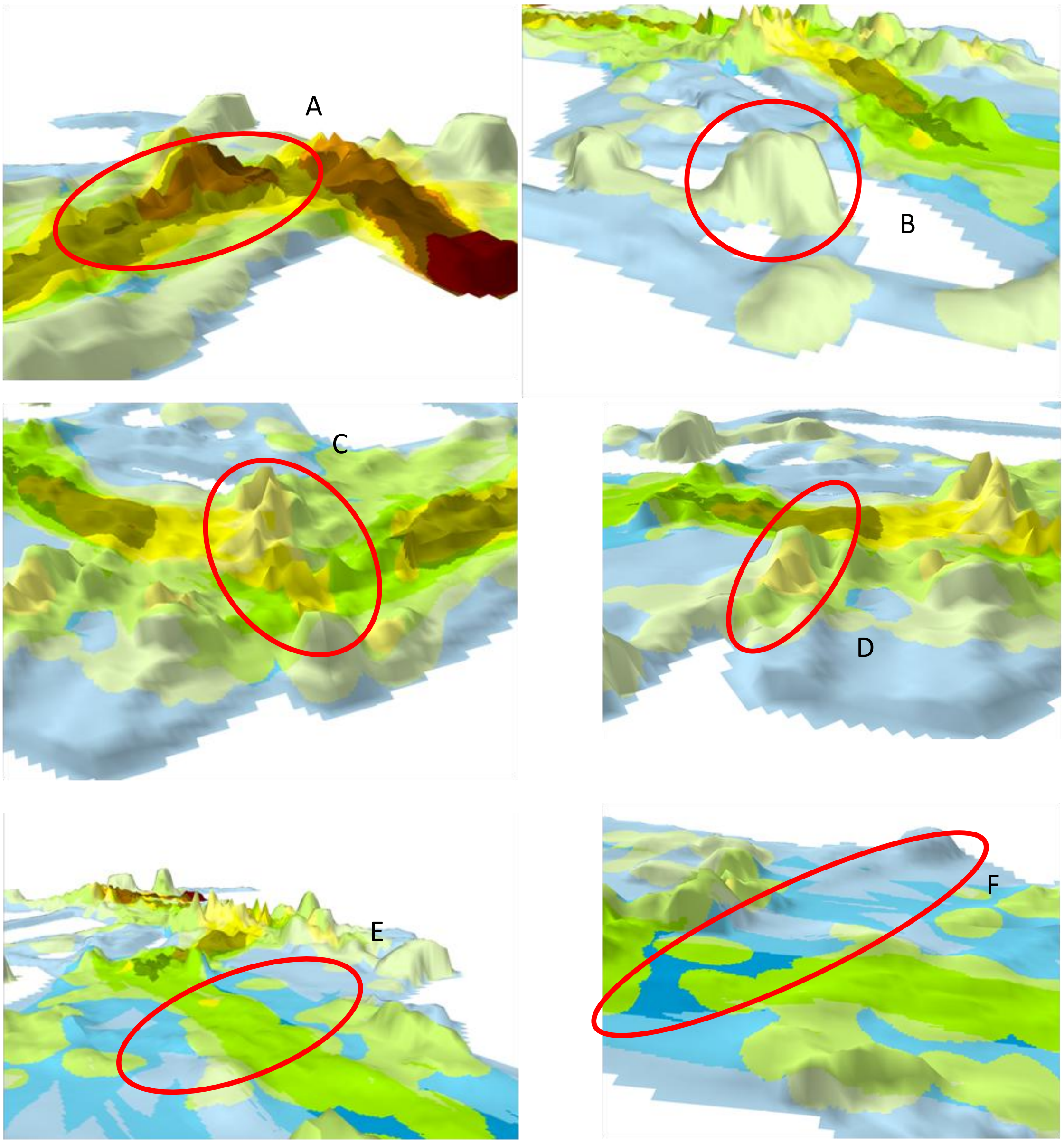


Figure 53: Data features: A) Data-mountain range, B) Data-peak, C) Data-ridge, D) Data-trench, E) Data-plateaux, F) Data-valley

7.5.2 Areas of Interest

Using the aforementioned special features, areas of interest within the data can be identified. By comparing these regions with features found in the environment we can draw conclusions as to what may be causing them. Unlike the earlier comparisons where individual datasets were used to identify the areas of interest, it is possible from this single visualisation to identify areas that are of interest in all datasets.

One area of interest is the region containing a 'blue data-valley' and the regions around it (Figure 54). The data-valley itself indicates a faster moving region with a relatively low track count. However, to the left of the data-valley we see a data-ridge form indicating that for some reason after travelling at faster speeds users tend to slow down. Near the north end of the data-valley we can see a data-peak formation indicating that at this point one or more users paused for a period of time to orientate themselves. The right of the data-valley is formed of another data-ridge, although this one is less pronounced than the one to the left of the data-valley resulting from again some slower average movements.

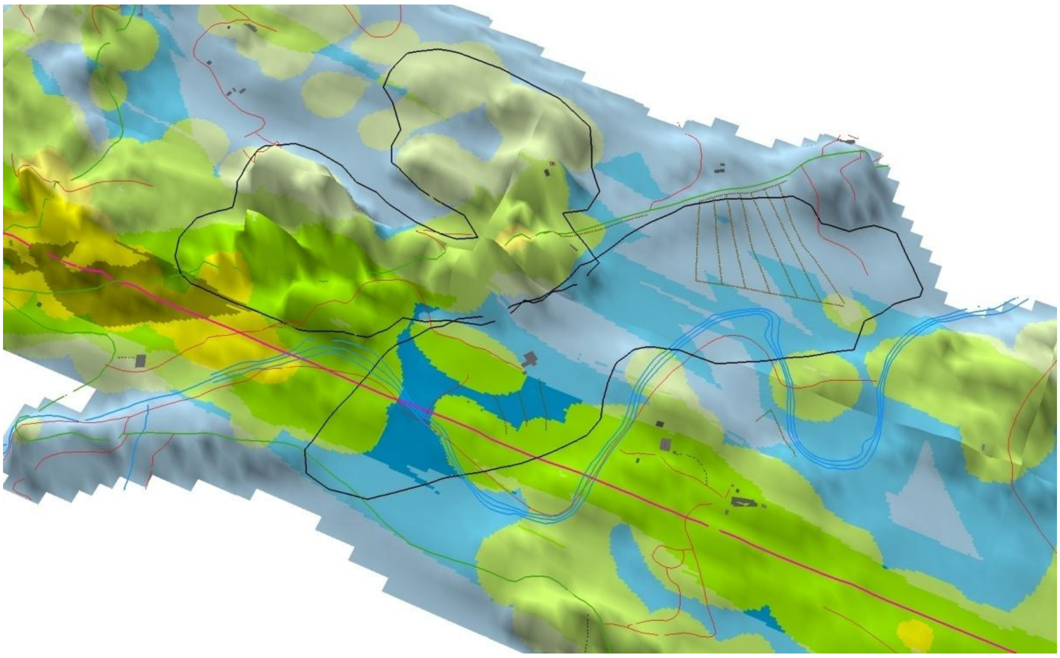


Figure 54: Blue data-valley and data-ridge area of interest

When we compare features in the environment and the data-valley and data-ridge features we can begin to understand why the fast movement occurs followed by a pause. We see that the movement begins after the river is crossed and ends upon reaching the main road and hills thereafter (see Figure 55). When looking at the intended route, the user can see that they need to cross the river three times before reaching the main road, and so when they reach it after only one or two crossings they become aware that they have gone along the wrong track and then need to make a decision as to what to do next. There are three main options at this point: carry on and try to head towards the end point by going in an estimated direction; retrace their steps and then try to go along the correct route; or follow the main road down to the point where it would cross the intended route and carry on from there. We can see that from the more intense shading and the data-trench going through the data-ridge that a number of users followed the road to the crossing point travelling at higher speeds.

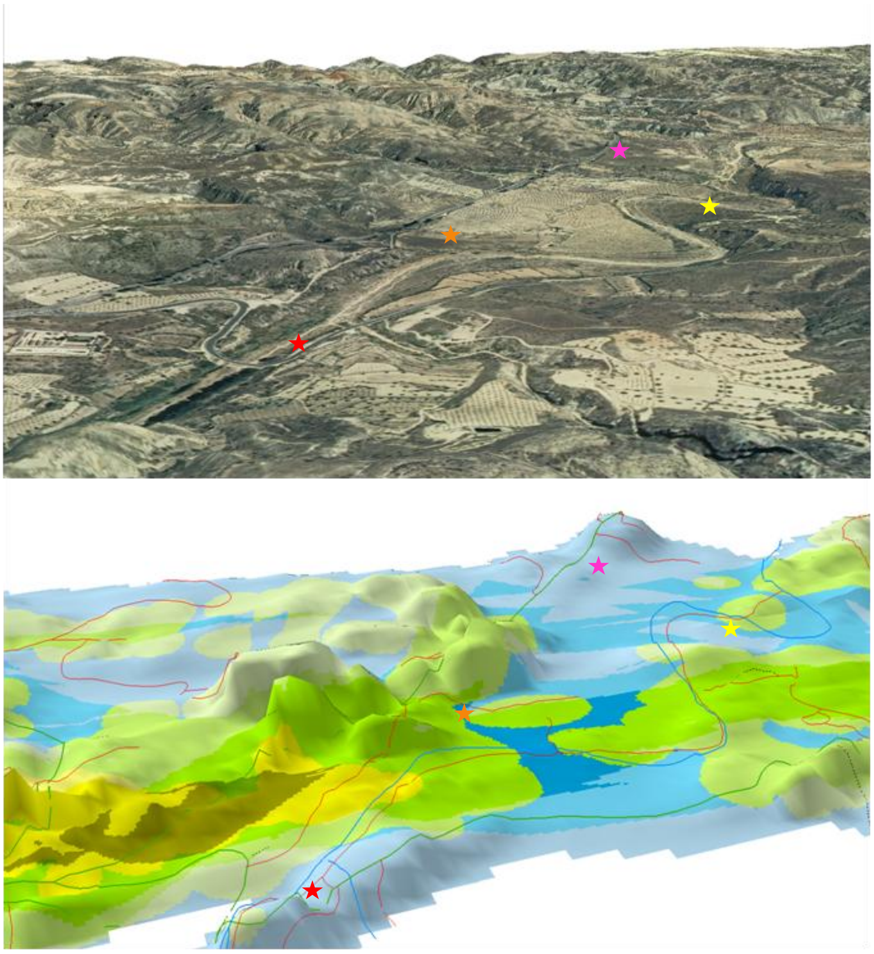


Figure 55: Blue data-valley comparison (coloured stars match locations in each environment)

Another interesting area is a data-ridge which appears just before the third turning point (Figure 56). The interesting aspect of this is that to the left of the data-ridge is fairly flat indicating a more constant speed whereas before the turn the data-terrain is much bumpier and darker. This implies that before the turning point a large number of users were travelling in the same area although there were pauses to determine their orientation. The darkness in the region resulting from a larger number of users would suggest some characteristic in the environment that draws users in and allows them to travel in a similar direction for a period of time. Again by superimposing the features in the environment over the visualisation we can draw some reasoning as to why these characteristics appear. Between the second and third turns where a large number of tracks are present there is a particular and distinctive feature where the higher density of tracks begins – the intended route intersecting a t-junction between two main roads. The reason for a number of users arriving at this location is that in many cases one of the two main roads was being followed. In some instances users overshot the second turn and carried on until they met with the road coming from the south-west, whereas some others came across the continuing road from the south-east and then carried on along. In fact, the road from the south-west forms a data-trench within a data-ridge which appears owing to users coming across the road when they were not expecting and then travelling along it to the intersection. At the end of the section, the bumps within the data visualisation in the area signify locations where the users paused as they approached the turning point. Again when comparing with the environment we see that this corresponds to a bend in the road that appears at the point of turning. The change in both clustering and track count indicates that this feature was used as a marker for turning, although in some cases either the turn was not taken and the user carried on, or they turned but to an incorrect bearing. It appears that in most cases users turned to the correct route (a 3rd level for the track count classification) and the majority of those who did not continued in approximately the same direction as they were travelling (a 2nd level for the track count classification).

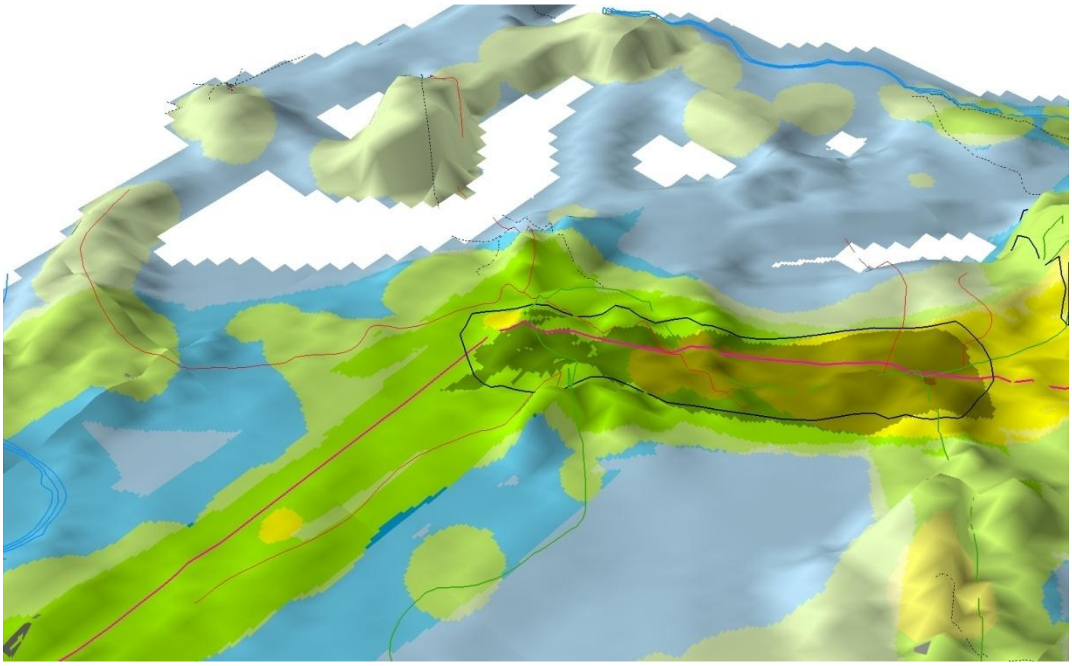


Figure 56: High congestion and track count area of interest

Just before the second turn we also see a data-ridge appear, although this one is more prominent (Figure 57). The north of this data-ridge may be attributed to the winding road passing through the area, where if people travelled along this road their points would become more clustered owing to the doubling back caused by the layout of the road. The southern portion is however not in line with any distinctive features in the environment and so is likely the result of users pausing so that they can determine their location. The colouring signifying the track counts and congestion indicate that in this region each are mid-low value with both ranging from classification 2 to 3. This indicates that although the clustering occurs it is more of an individual aspect meaning that users would not all be travelling slowly along a linear feature but generally stopping somewhere which happens to be in a line with each other.

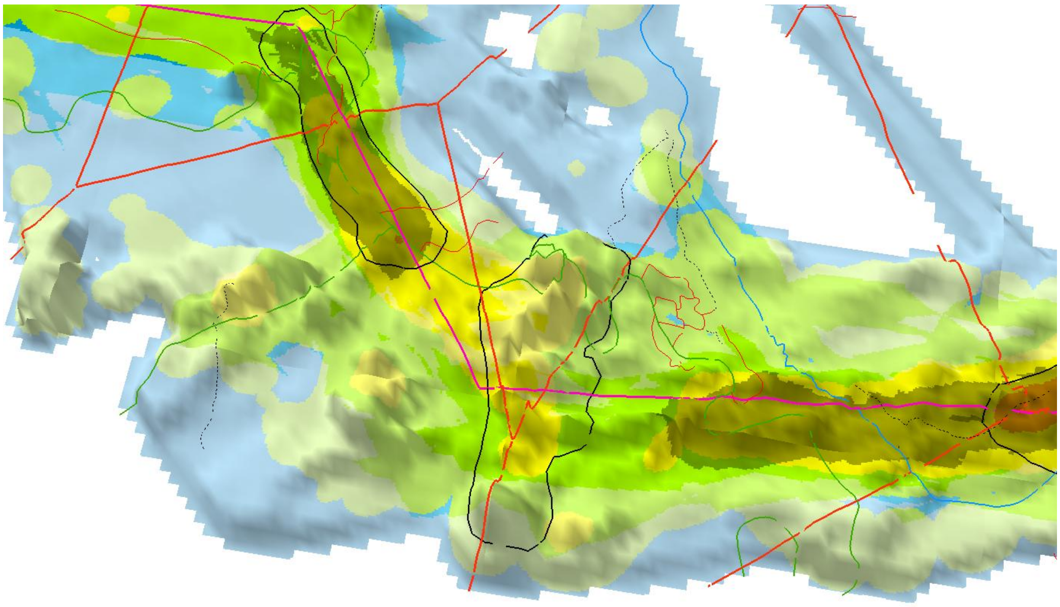


Figure 57: Pre-turn data-ridge area of interest

When compared with environmental features however there are no obvious linear objects along this data-ridge feature. The possible answer to why this data-ridge occurs in such an apparent linear formation may lie with the info-marks implemented in the experiment. When the transition boundaries are displayed it can be seen that the southern peaks in the data-ridge coincide with the boundary between two info-mark features and that the northern ones fall within the southern region of a different info-mark region. This suggests that the ridge may be formed by the info-marks as opposed to any physical feature within the environment. When we view the clustering datasets for each info-mark delivery method however, we in fact see that that was most probably not the case as there is no clustering occurring along the boundary for the PDA delivery method, but some data-peaks appearing along the boundary where the info-marks were not actually presented to the user (Figure 58). Therefore the reason for the data-ridge data feature may be due to users being able to successfully implement a 'dead reckoning' orientation method to determine the location of the second turning point.

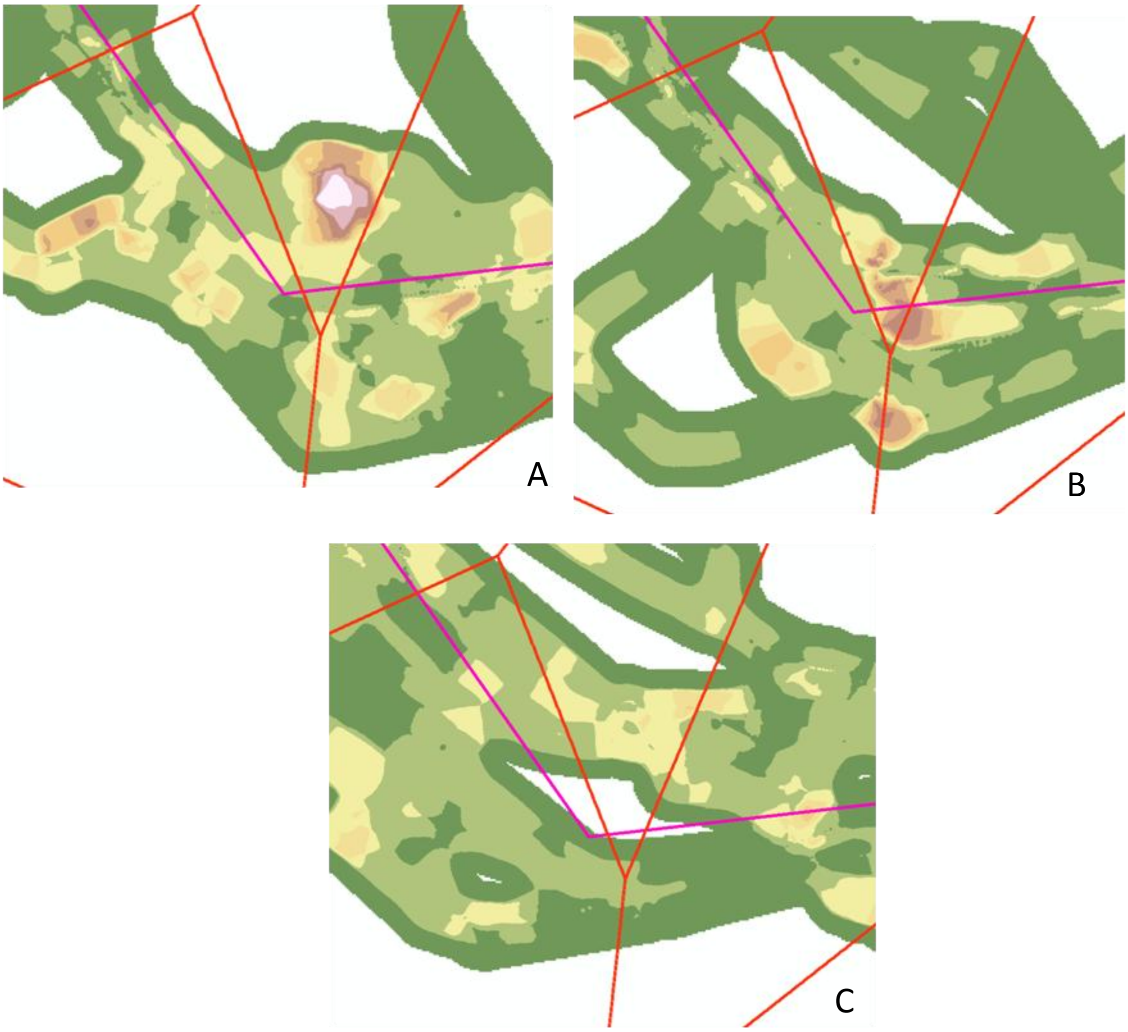


Figure 58: Clustering between info-mark delivery methods A) no info-mark. B) onscreen info-marks, C) PDA info-marks

The final area of interest is located just after the first turning where a darker region forms of brown colouring (Figure 59). Before the turning a large region of similar characteristics is present, but this is not particularly surprising as along that section the users simply had to face the correct compass bearing and walk. The interesting aspect of the region after the turning is that users slowed down or paused before as a whole altering course slightly to follow the intended route more accurately. Upon investigation, we see that this region coincides with an intersection between the freeway and river features in the environment which occurs along the intended route. This means that some people miss the first turning or head in the wrong direction but then manage to correct themselves based on this feature.

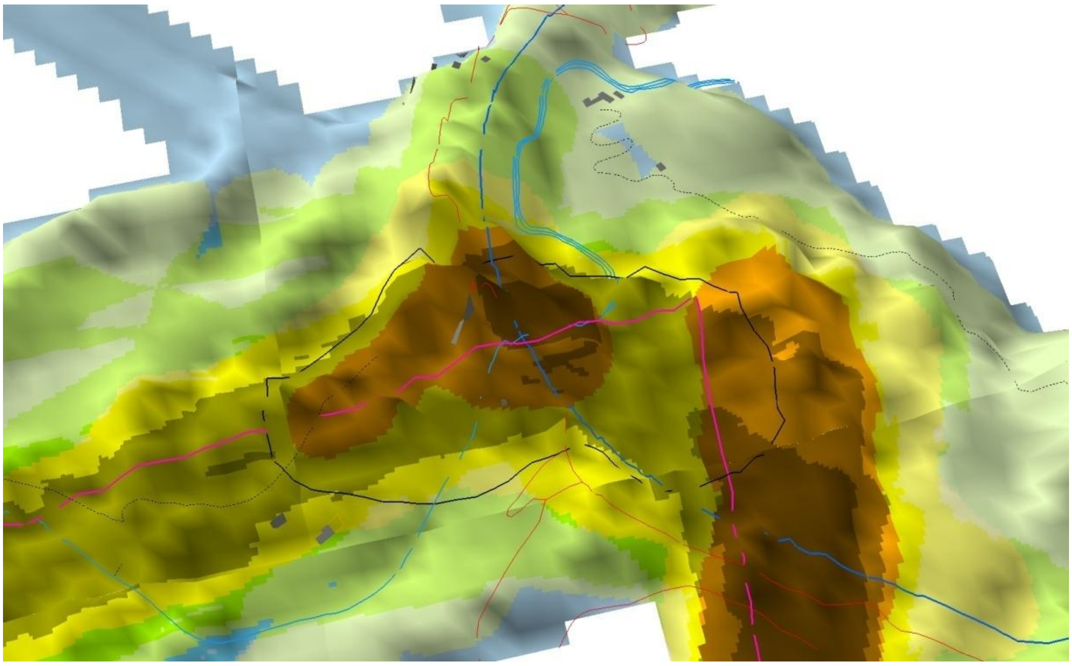


Figure 59: First turning point area of interest

7.6 Data Visualisation Against the VE

As a further means of analysing the data the areas of interest identified were compared to the same location in the VE. Although comparisons had already been made between features found in the environment using the GIS representations, this was not necessarily what the users saw whilst travelling. For example, without knowing the terrain or land cover classifications we could not say at what point the user would be able to see the feature that may be having an effect on the tracks.

Therefore viewing the area of interest as the user would have seen it is beneficial. To do this a Blueberry model was generated using the extracted data as used in the previous visualisations with the Blueberry functionality of elevation and orthographic imagery used to project the data. The same elevation data (the level of clustering) was used by exporting the DEM as a tiff image for importing into the Blueberry tiler tool. The shaded regions identifying the level of congestion and physical features in the environment were exported from ArcMap as a high resolution tiff image. To allow for better distinguishing between peaks within Blueberry a hillshade raster was generated from the clustering raster using the ArcMap spatial analyst tool (shadowing generated by Blueberry itself was not satisfactory). This was then displayed with a high transparency (70%) to ensure that the shading did not alter the colouring used on the congestion/track count raster excessively. The map generated for use as the orthographic photo can be seen in Figure 60 (the resolution and size of the features are different between the map presented and the one used in Blueberry to allow for suitable display of the information).

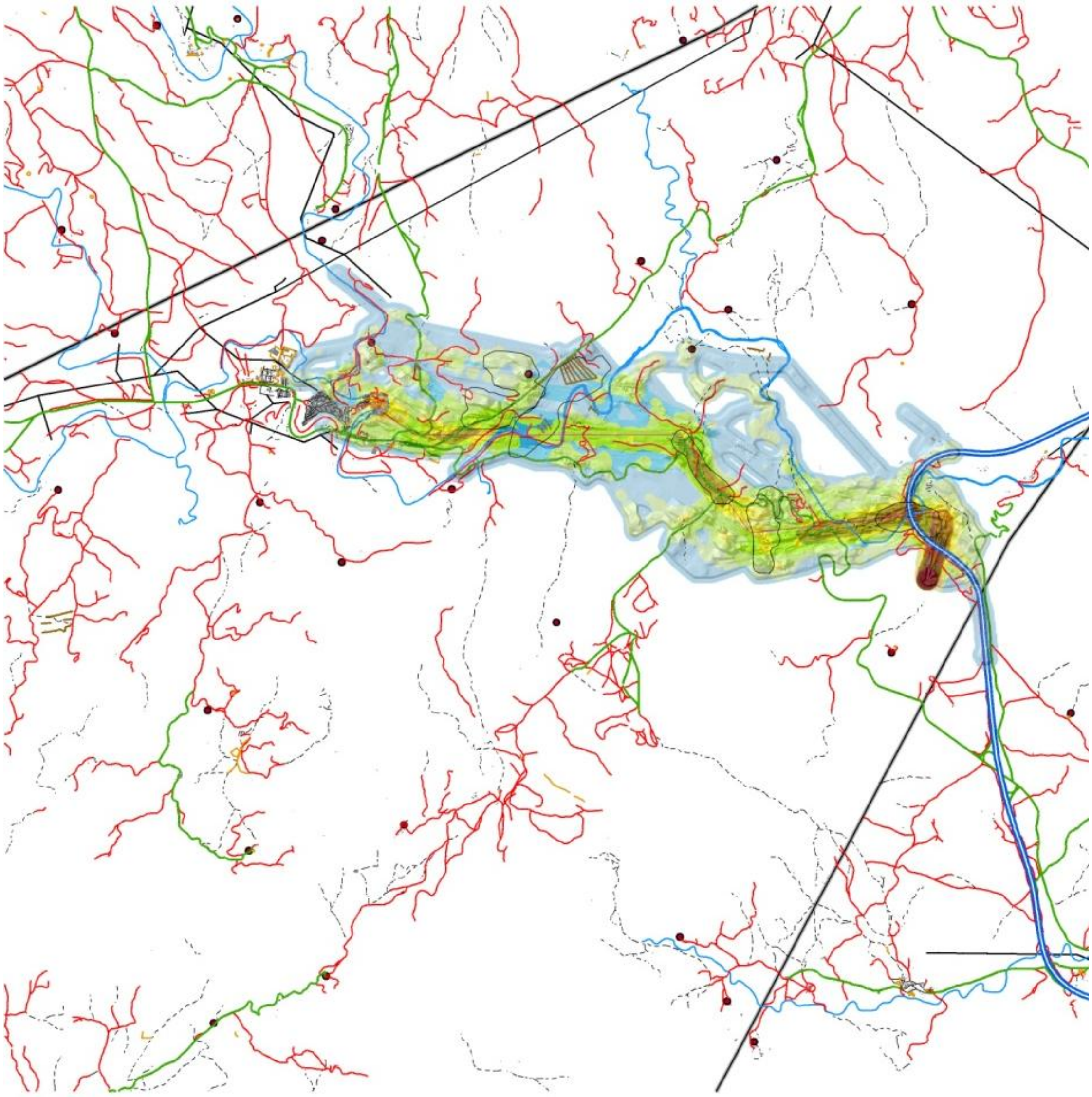


Figure 60: Data visualisation and environment features

Initial attempts to compare the data side by side with the VE using the VegaPrime system were made, whereby movement in one scene was generated from the movement in the other. By setting one observer to cover half of the screen and the other to cover the other half and updating both scenes from the same motion model it would be possible to achieve this visualisation method. However, owing to licensing restrictions only one Blueberry scene could be displayed at a time. Therefore the views from the same location were obtained by determining the location of the view in the data visualisation (a key press of ‘p’ to give x, y, z coordinates as well as observer rotations) and then applying the same observer location to the real environment. An offset in elevation to account for the different elevations present in the two environments was also used. The boundary regions signifying the areas of interest were displayed in the environment itself by adding an extra vector representation to the model with the data source being the area of interest outline shapefile.

7.6.1 Blue Data-Valley

When we look at the 'blue data-valley' region identified in the data formed by users moving quickly across the terrain, we observe particular aspects of the environment that could cause this (Figure 61). Firstly we see that the terrain is particularly flat (green arrow) meaning that users can see features that could become landmarks in the distance. This means that they can simply point towards these and then travel in a straight line. Secondly we see the area with the most track counts (the darker blue and green region) corresponds with the area of the intended route that moves close to the river and its neighbouring path (red arrow). In the environment this feature is clearly visible both at a distance and up close and can serve as a fairly good landmark feature as the user knows they need to cross it three times in a straight line which can be planned from a distance.

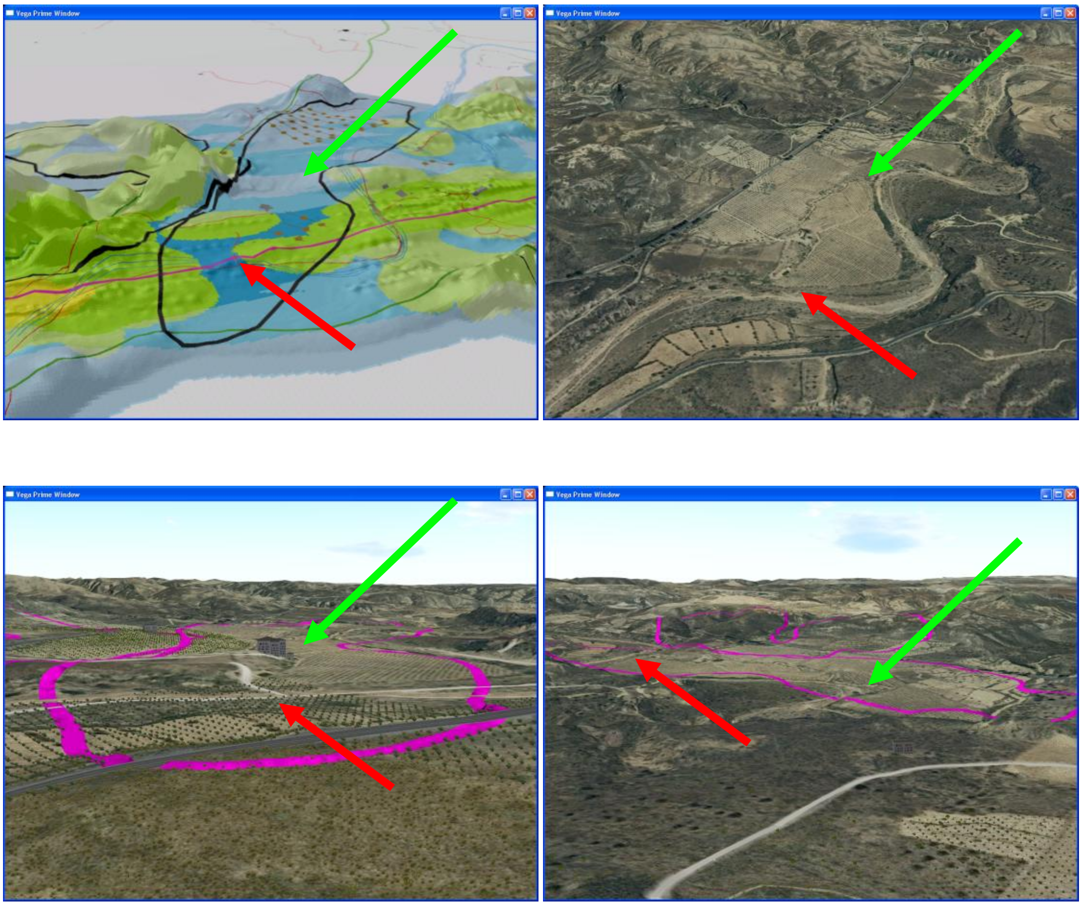


Figure 61: Blue valley environment comparison

7.6.2 Data-Ridge and Data-Mountains

The second area of interest is to the right of the blue data-valley region and is a data-mountain range feature (Figure 62). The first thing that was noted about this area of interest was the data-ridge formed by the data-mountain range with a data-trench generated by users travelling along the road (red arrow). We also see that past this data-trench the data-mountain range continues (green arrow) indicating that some people continued past the road before pausing. When comparing this to the actual environment we see that this data-mountain range coincides with a higher elevated region. We also see that the edge of the interest area falls at the peak of the hill (pink arrow) indicating that users were making use of this elevated region to determine where they were located.

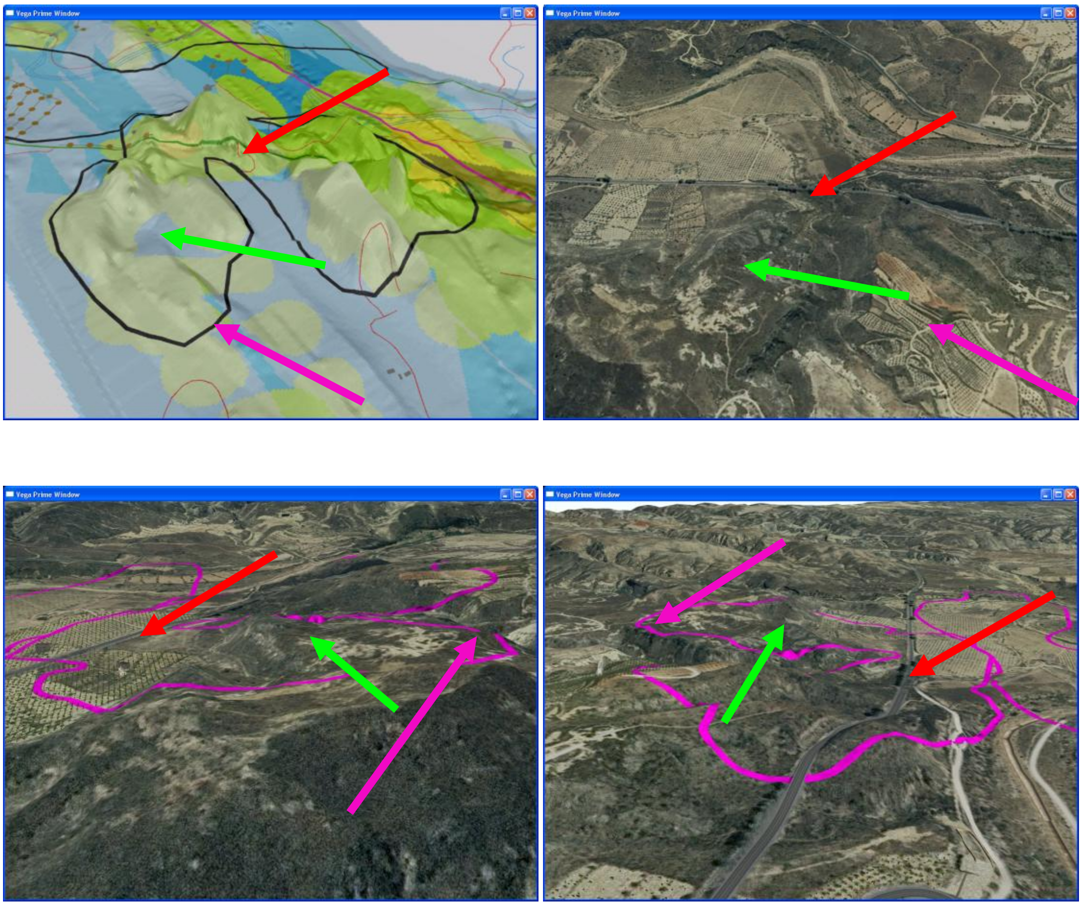


Figure 62: Ridge and mountains environment comparison

7.6.3 Track Convergence

The area of interest within the centre of the tracks is interesting owing to people converging on the same place at relatively the same time within their journeys (Figure 63). It was identified from general analysis that this could be owing to the junction between two main roads present along the intended route (green arrow) and then the following of this road for some distance. Comparing the data visualisation with the actual environment we do not see any other physical features such as elevation changes that could account for this convergence of tracks. At the other end of the area of interest we also see the slightly peaked data region (red arrow) signifying a general pause by the users. As discussed this occurs before the turning and so is likely the result of users reaching a point where they feel they are needing to change direction which in this instance could be the large bend in the road.

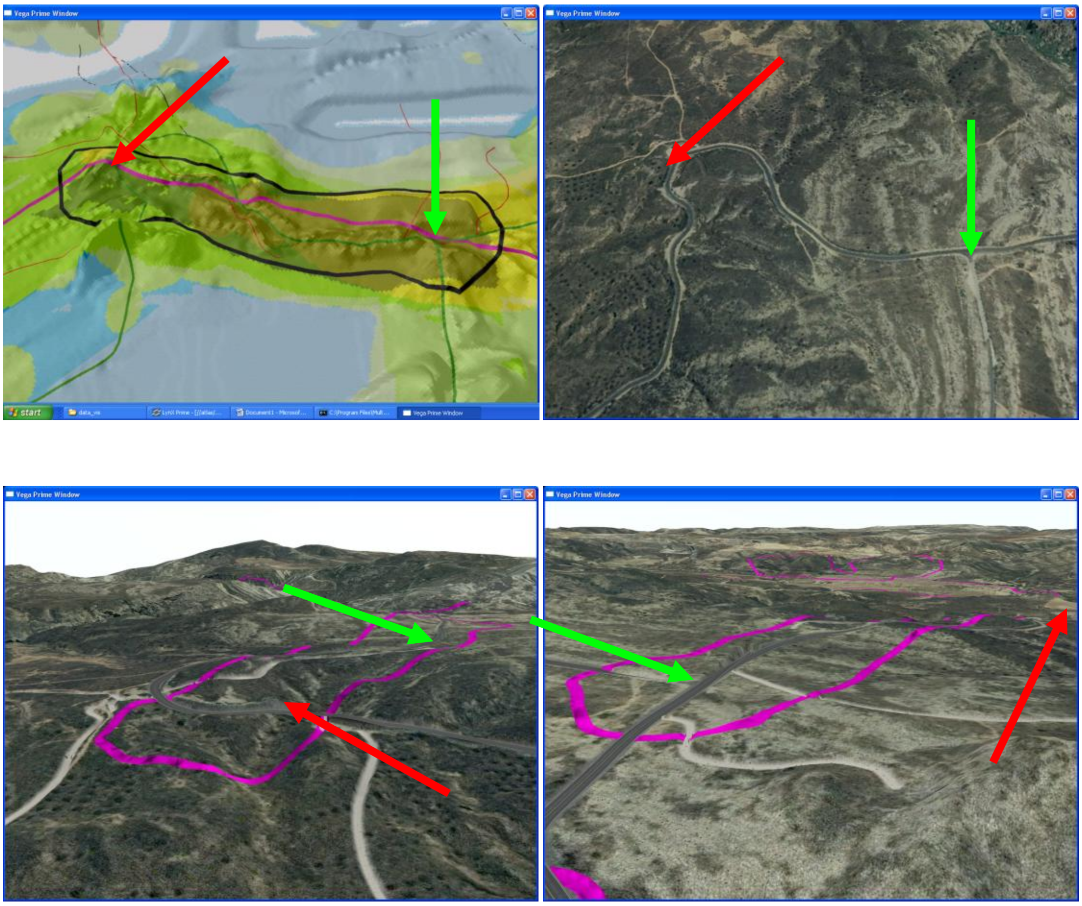


Figure 63: Track convergence environment comparison

7.6.4 Data-Mountain Wall

One of the more prominent mountainous data regions found was before the second turning (Figure 64). Just by looking at the data, one of the data-peak areas in this region could be owing to the winding road feature (pink arrow) whereby if people were travelling along this road the points used to derive the clustering data would end up close together due to the road doubling back on itself. However, the larger data-peak below the road corresponds with a peak in the environment (green arrow) indicating that users may be using this high ground to identify their location. A similar feature appears in the bottom of the region where again a data-peak forms on the top of a ridge in the environment (red arrow).

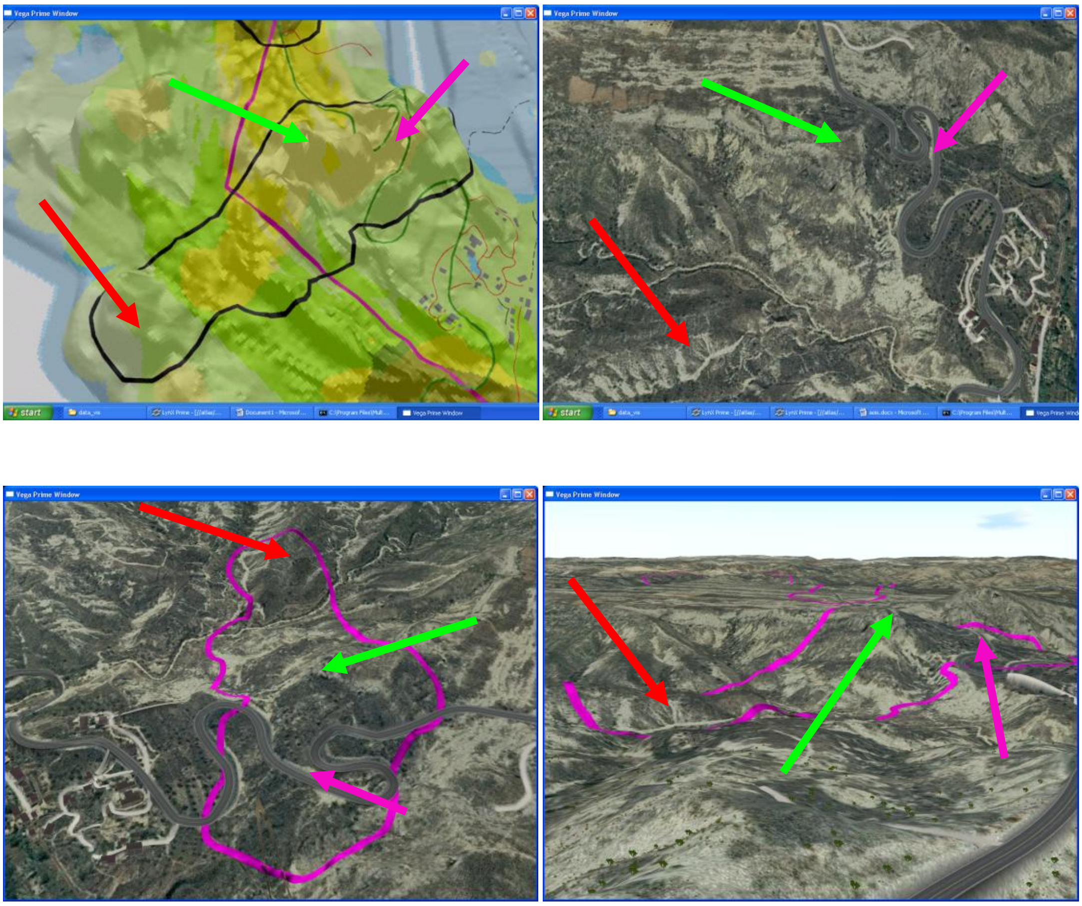


Figure 64: Mountain wall environment comparison

7.6.5 Dark Data-Mountains

The 'dark data-mountain' region is located near the start of the tracks and is characterised by a mountainous data region which is darkly coloured indicating that a number of users were travelling slowly with a large number of tracks being in the same location at relatively the same time (Figure 65). The largest data-peak in this area is just after the users cross the freeway which corresponds to a location in the environment which is north of some buildings and a minor road (red arrow). The pause here could be accounted for by users attempting to identify the distinctive building features in the map as a means of determining their location. The data-peak regions to the right of the freeway and intended route appear on the side of valley wall feature within the environment (green arrow) and can be accounted to people pausing on the slope to turn and look towards the river/freeway intersection and then identifying from their compass bearing whether they have travelled far enough.

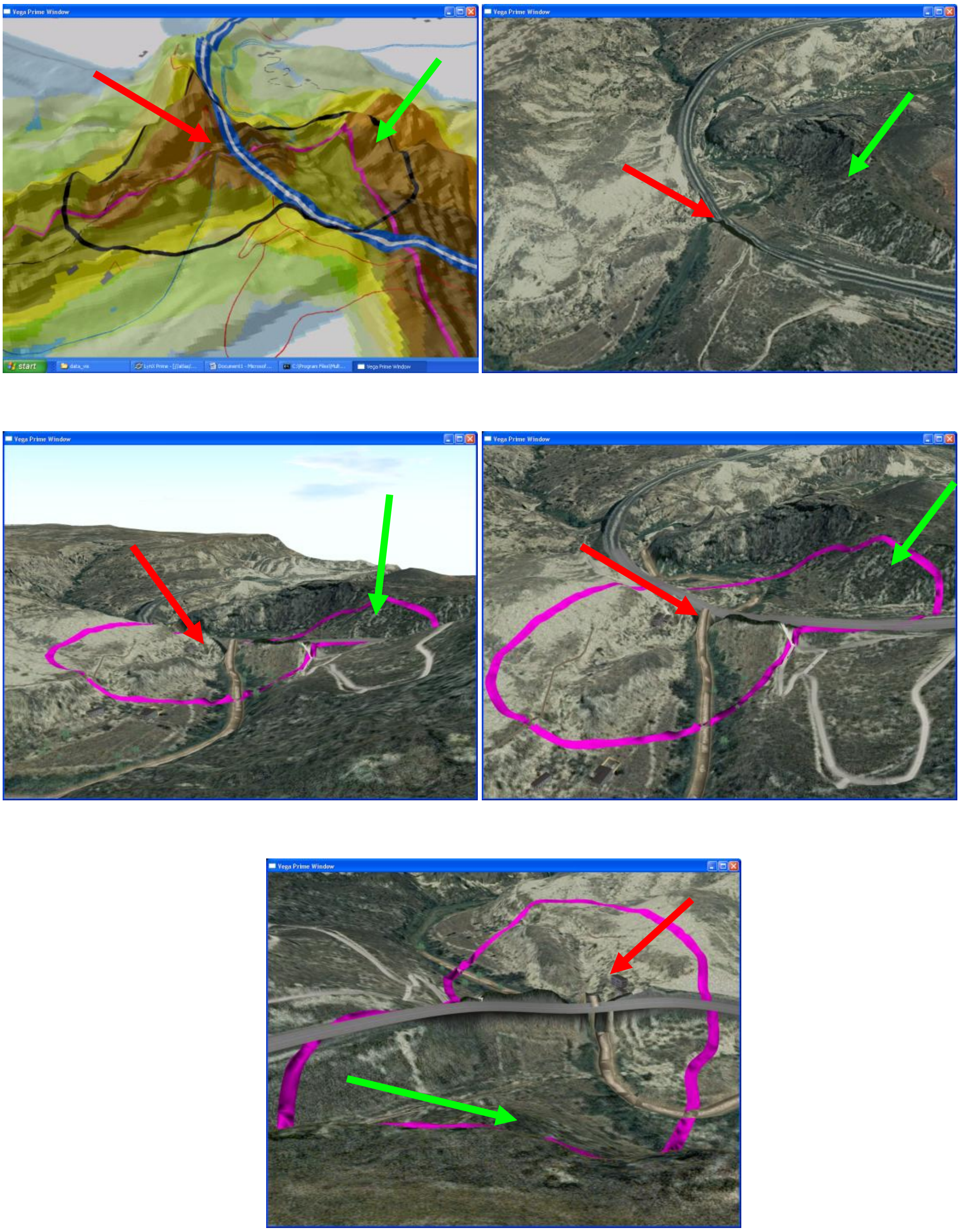


Figure 65: Dark mountains environment comparison

7.7 General Interpretation of Trial Results

From the visualisation implemented we see that users followed the first third of the route fairly accurately, being shown by the higher levels of congestion and track count. Some tracks did deviate by a large amount from the intended route but these were in the minority. The higher track counts continue past the first turning until diminishing at around three quarters of the way through the second portion of the route, with several tracks following the river and roads in the region. A data-ridge appears before the second turning as a result of people needing to pause to determine their orientation within the environment before making the turn. After the turn, the tracks begin to converge again resulting in a higher level of track counts and congestion owing to a key feature occurring along the intended route – a t-junction between two main roads. The convergence occurs owing to users following one of the two roads and thus ending up in the same place along the intended route. After this junction, users continue along the intended route until the third turning which is located just after a distinctive curve in a main road. At this point the darker shadings indicate that two main occurrences happen at this point: either the users carry on in the direction they are already heading, or they turn in roughly the correct direction of the route. Further along the route we see the blue data-valley formation indicating that after crossing the river the users travel at a faster speed with little or no pauses. If the users are travelling along the intended route, the width of this data-valley continues until the third intersection with the river. If on the other hand they have travelled along a route which is too far north then the data-valley stops at a data-ridge formation generated from the users approaching a road before expected. This causes them to re-orientate themselves and then travel along the road (causing a data-trench) or continue on the route they are taking. Around the end point we see multiple data-peak features caused by locations where individual users paused to determine their location. At the end location itself we see that there are a higher number of tracks with a higher congestion. This indicates that a number of users did finish their trial in the area of the intended location. The actual track count value at the final location is 15, meaning that there were 15 tracks within the vicinity at some point in time, although to look at the actual end points of tracks we find that 12 are within 100 metres of the intended end point.

The later stage of analysis was to compare the data visualisation with the actual environment visualisation. Several physical features that could contribute to the formation of data features were identified from a simple map overlaid onto the visualisation. The identification of further aspects that could cause the data features was achieved by comparing the data visualisation with the VE. One of the main characteristics highlighted was the often correlation between data peaks and physical peaks within the environment. This relationship shows that when travelling in the environment people were using the high ground as a means of determining their location. Another area where the terrain appears to have an effect on the navigation was in the 'blue data-valley' region caused by fast movement through the environment. When looking at the actual environment we see that this region is particularly flat and so it is possible to see features in the distance. This means that users can simply identify a target to move towards and then travel in a straight line in the direction of that feature.

The main benefit of these types of visual analysis is that they can aid in the development of future environments. For example the most efficient environment would consist of just a dark data-plateaux indicating that people moved quickly without pausing to determine their location. As seen from this study this type of feature forms when the user is able to see a target destination and then head in that direction (in the case of the data collected here the data feature generated was the blue valley). If turning points are required then it has been seen that people tend to use peaks in the environment as a stopping point to identify their location and trajectory when other features are not available. Therefore it would be beneficial to have such peaks at turning points. Finally it can be seen that roads in the environment form troughs in data mountain ranges. These occur due to users travelling along the roads to get back onto the intended route. Therefore in a new environment there should be distinguishable roads at intervals that would lead the user back on track in case they deviate.

Many of these characteristics link with the UIT feature types. For example, the data has shown that data-trenches coincide with routes multiple users travel along after making a decision, synonymous with the path classification. Physical peaks in the environment correspond with locations where users are making decisions which links with the node classification. Finally flat areas in the data signify a distinct and constant movement which can correspond to districts owing to the area in the environment having some characteristic which makes navigation there easier. Edges are less well defined in the data as there were no real edge features in the environment that created a physical barrier to movements. If restrictions were placed on movement at edge forming features (i.e. rivers and steep slopes) then this would likely be found in the data as a peaked area with a number of track logs on one side of the feature followed by no peaks and no track counts on the other side of it (as users would not travel into it). Landmarks are also more difficult to observe in the data due to how they are interacted with. There are two possibilities when a landmark is used in navigation as discussed earlier. They can be features used to confirm that the correct route is being travelled, or features signifying a particular turning point. How the landmark was used would alter how it would appear in the data, as in the first case it would most likely be travelled past without stopping, but in the second case it would likely be that the user would stop to determine the direction to travel in. Therefore, when used as a correct route cue it would be difficult to identify in the data as it would not be associated with any data-peaks, although a small amount of data-peaking may occur if users slowed down as they approached to confirm that it was the correct landmark. There may be an increase in track count at the location but only if the landmark was visible from a distance so that users could travel in its direction. In the turning point instance, it is likely that the landmark would appear as a data-peak in the data similar to those found at peaks in the environment. Therefore, if there was a peak occurring in the data which corresponded with a feature in the environment it is likely that that feature is being used as a turning point landmark (in effect, the landmark becomes a node).

Most of the items just mentioned with regards to improving navigation in a newly created environment revolve around the insertion of features. Such alteration of the features present would then result in the environment becoming unrealistic in comparison to a real world location that it may be modelling. However, we can use this analysis to suggest possible improvements to a virtual representation of a real world place in terms of either mobile assistance systems or insertion of realistic aids. For example, here it has shown that peaks are used as high ground for determining location and so therefore it may be beneficial to add signposts at the top of each peak highlighting the direction to specific locations. Roads were found to be used as channels so a number of signposts placed on the road identifying where it is heading to may be useful. Finally pauses happen in the environment when people look around. In the real world this may not occur as often as you can look whilst still travelling in a different direction and so the actual interaction methods with the environment could be adjusted to allow for this (i.e. travel in the direction you are facing, not the direction you are looking). This would then allow users to continue on their path whilst looking around and so may reduce the number of stopping points.

7.8 Dataset Comparison

As well as using the visualisation method to compare the data with the environment it can also be used to investigate differences between datasets. In the route tracing exercise the track logs generated from users were either presented with the info-marks onscreen, via a PDA or not at all. Therefore we can use the visualisation technique to investigate whether any differences between the datasets can be visually identified.

The same steps were taken to create the visualisation as in the full data visualisation except that the tracks were split into their respective groups thus resulting in slightly different values being used as the classification boundaries using the Jenks Natural Breaks method which can be seen in Table 35.

Track Count			
	Info-mark display method		
Classification	None	Screen	PDA
1	2	2	2
2	4	4	4
3	6	6	6
4	8	8	8
5	10	10	10
Congestion			
1	317.536	331.008	359.019
2	621.150	607.057	765.596
3	1091.752	1101.040	1388.167
4	2169.581	2176.178	2201.321
5	3915.361	3759.827	3255.880

Table 35: Jenks classification boundaries

By comparing the visualisations for the track logs from different info-mark delivery method we can observe some differences between the data. Firstly, one of the more striking differences is with regard to the road junction near the middle of the route where in the full dataset it appeared that many users were converging on it at the same relative time in their tracks (see section 7.6.3). When no info-marks were presented we see that a number of users arrived at this location (congestion = 4, track count = 3) at relatively the same time (Figure 66a). In the onscreen delivery dataset we also see that a number of tracks meet at this location (congestion equal to four, track count equal to four) although this is a smaller region before congestion drops to three with the track count remaining at four (Figure 66b). This signifies that when the info-marks were presented via the on-screen delivery method a large number of users travelled through this region and that they arrived at the junction at relatively the same time with users then travelling at different speeds as signified by the drop in congestion. However, when we look at the PDA delivery method dataset we see that there is no such congestion or track count values and that the tracks are actually more spread out signified by the data-plateaux region (Figure 66c). In fact, at the junction we only see values of congestion equalling two and track count also equalling two, signifying that when the info-marks were presented to the user via the PDA they did not visit the road junction even though it was along the intended route. One reason for this could be that the users only used the info-mark information to determine their path, so as long as they were within the area of the junction this was perceived as satisfactory in regards to accuracy. We also see far less data-peaks indicating that users did not pause to determine their location as they did in the on-screen and no info-mark visualisations.

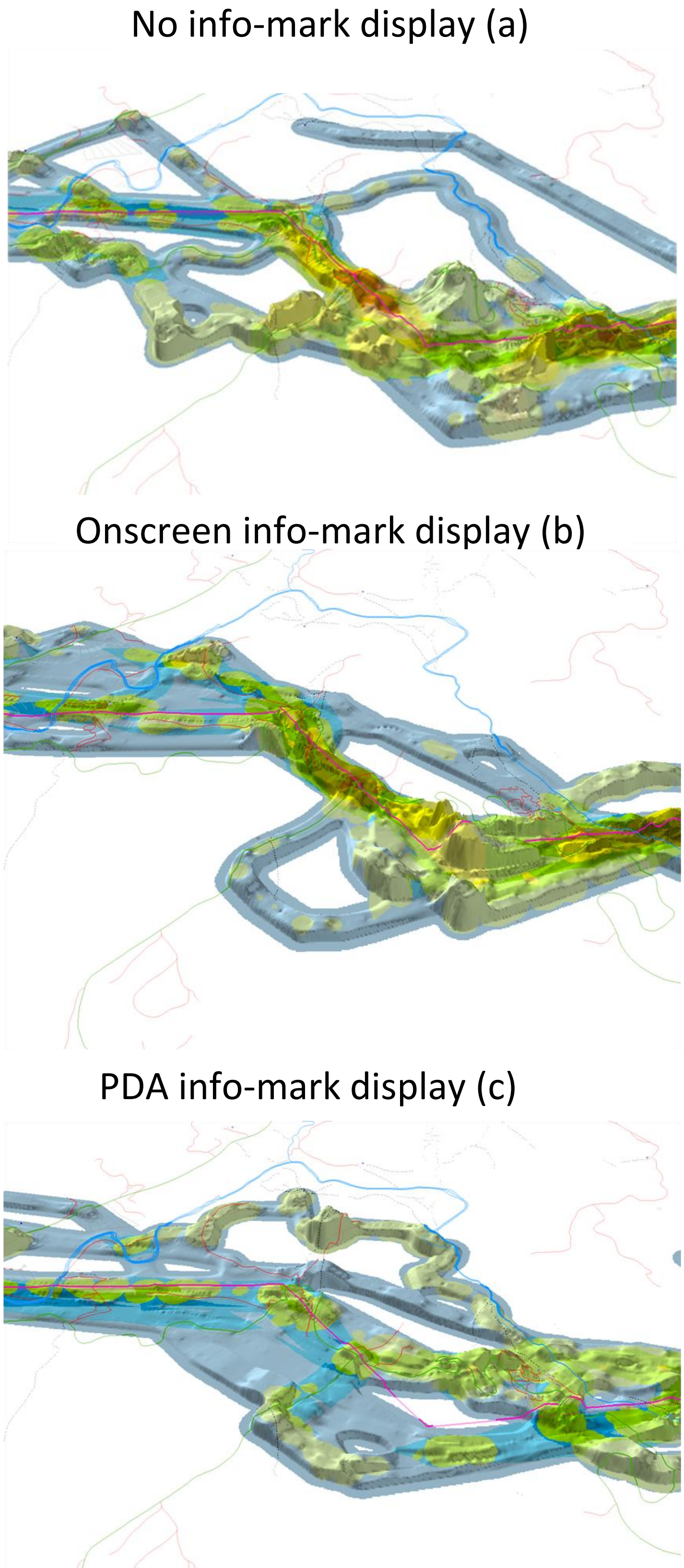
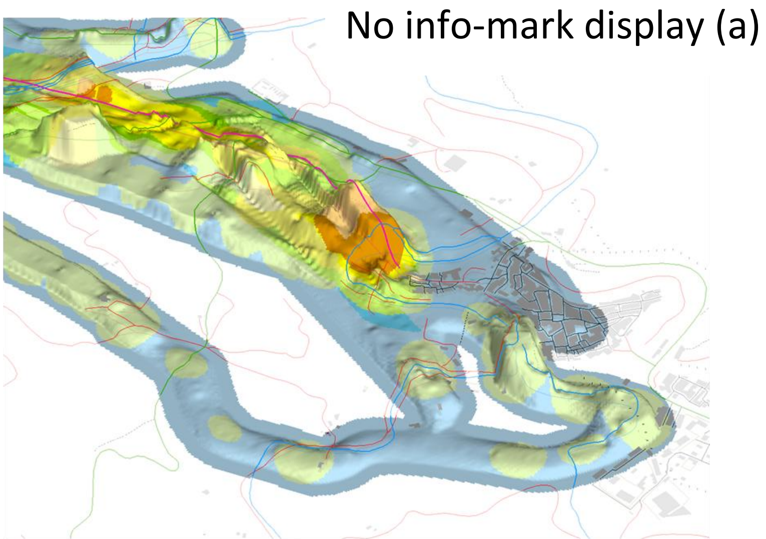
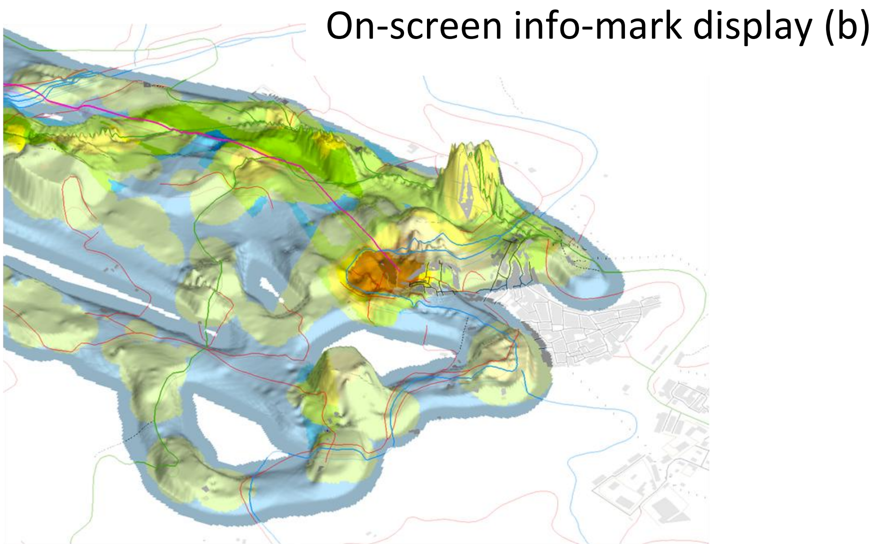


Figure 66: Central region info-mark delivery method comparison a) No info-marks; b) onscreen info-marks; c) PDA info-marks

Another interesting location when comparing the datasets is the end of the route. In both the on-screen and no display datasets (Figure 67a & b) we see a dark region close to the end point indicating larger numbers of track logs and higher levels of congestion (congestion equal to foue and track count equal to three to four for on-screen display, and congestion equalling three to four and track count equalling three for no display). However, for the dataset where the info-marks were presented via the PDA (Figure 67c) we see lower values for both congestion and track counts (congestion equalling one, track count equalling two) indicating that fewer users actually stopped at the intended end location. Again this could be accounted to the users relying only on the info-marks and so that when an info-mark appeared which corresponded to a location near the end point users stopped then rather than actually attempting to find the exact location.





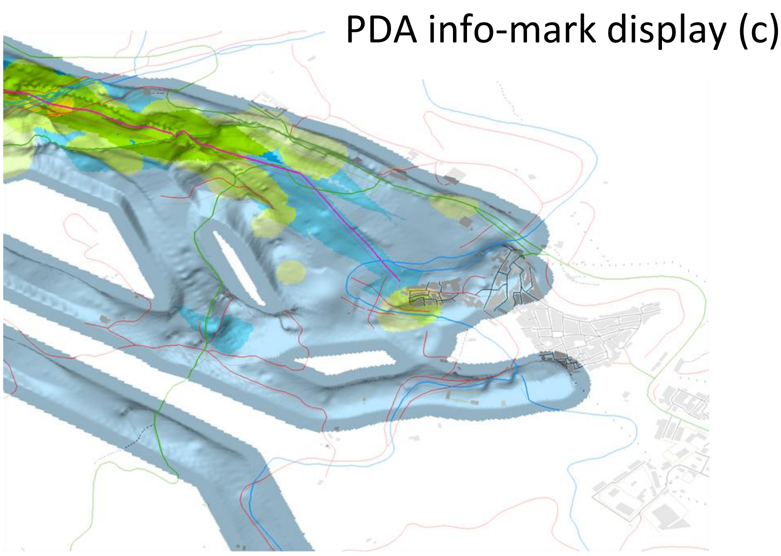


Figure 67: End point region info-mark delivery method comparison; a) No info-marks; b) onscreen info-marks; c) PDA info-marks

Near the beginning of the route all datasets look fairly similar (Figure 68), showing high numbers of tracks heading to the north east before turning close to the intended turning point. In all of the visualisations we see a somewhat bumpy region before this turning point which can be accounted to people pausing to line themselves up with the intersection of the river and freeway with a matching compass direction of the intended route.

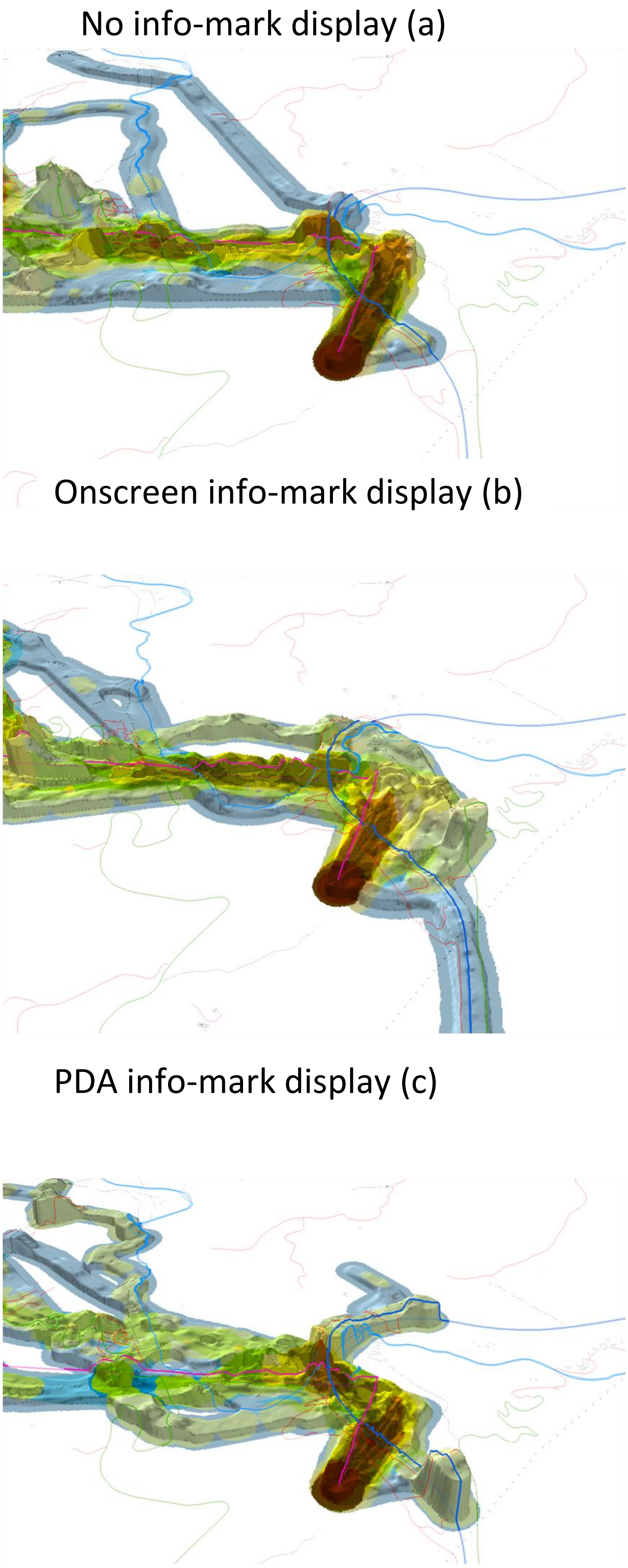


Figure 68: Start region info-mark delivery method comparison; a) no info-marks; b) onscreen info-marks; c) PDA info-marks

When we look at the dataset visualisations as a whole (Figure 69) we see that although the PDA delivery method dataset appears more spread out indicating that users did not follow the route as accurately, we do see that it also appears flatter than the other two visualisations indicating that there were fewer pauses and less slower moving sections. Up until the second turning point along the route (travelling from the right) the visualisations do look fairly similar except for the parts signifying where individual users ventured a large amount of distance from the intended route. In all the visualisations we see a higher number of users travelling roughly along the intended route indicated by the darker shading of the colours. The second turning however is where the datasets begin to look drastically different with the PDA dataset appearing more spread out than the other two. This could be owing to users relying more on the areal aspect of the info-marks after this turning (when presented via the PDA) rather than any distinctive landmarks in the environment. This would mean that as long as the users were travelling between info-mark regions in the correct order they would feel that they were travelling along the correct route. This can be seen at the intended end point as well, where lower levels of both track count and congestion are present in the PDA dataset compared to the other two, possibly as a result of users having the info-mark close to the end mark appearing and then deciding this meant they were close enough to the end point.

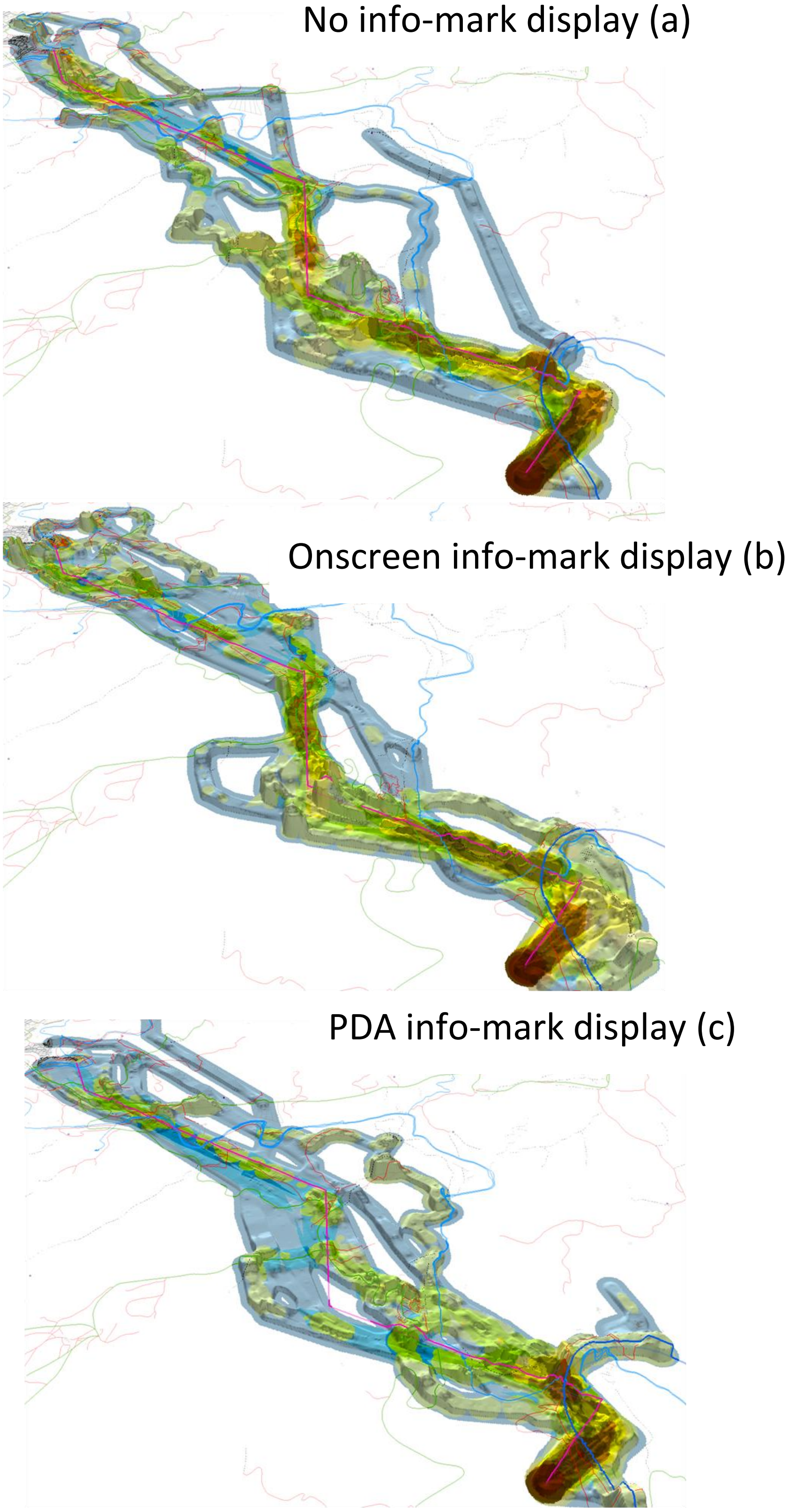


Figure 69: Global view info-mark delivery method comparison; a) no info-marks; b) onscreen info-marks; c) PDA info-marks

Overall, although the individual visualisations look relatively similar in shape we do see a difference in the PDA dataset from the other two. This seems to indicate that a different method was used to navigate the environment when the info-marks were presented via the PDA than was used in the other two. This in turn suggests that when the info-marks were presented onscreen they were not used as much by the people attempting to follow the intended route.

7.9 Summary and Advancements

The aim of this study was to identify a method of visually analysing the collected track log data so that conclusions about performance and user preferences could be derived. Here a method of combining three pieces of useful track log data (track count, clustering and congestion) into a single visualisation has been proposed and used to identify characteristics of the data caused by features within the environment. These characteristics have been defined using geographical features (data-mountain ranges, data-valleys etc.) to easily identify areas of interest and corresponding environmental features. It has been found that many peaks in the data formed by higher levels of clustering coincide with peaks in the environment as well as areas immediately before route turning points. Similarly flat areas in the clustering data tend to correspond with flat areas in the environment itself. Additionally, darker areas caused by higher track counts and congestion have been found to occur in areas of the environment where intersections of key features fall within close proximity of the intended route. Overall this study has identified that visualising multiple aspects of data in a three-dimensional way can provide insight into user performance. Although many of the data features identified are dominated by one of the pieces of data being visualised (normally the clustering depicted by elevation) others draw from more than one of the datasets. For example the ‘blue valley’ region is largely of interest owing to the flatness created by the low level of clustering, but the ‘blueness’ indicates as well that people were generally there at different times in their tracks. The end location of the intended route however is dominated more by the level of congestion and track count as opposed to the elevation provided by the clustering data. This signifies that a number of users visited this location (the shade) and that the congestion was high (brown colour) meaning they were there at relatively the same time. This can then be interpreted as meaning that a number of users ended their trials at this location with little amount of time spent confirming their location.

As well as viewing all the track logs in a single visualisation to derive information about the environment and how users navigated, different datasets were analysed depending on which info-mark delivery method was used during their trial. With this comparison it was clear that differences were present between the PDA delivery and the other methods. When visualised, the PDA dataset was visibly flatter and more dispersed than the other two suggesting that although users tended to pause less, they did deviate more from the intended route. We also see how these differences occur in specific places with the beginning sections looking similar in all datasets as well as the region heading west after the third turning point (travelling right to left), with the end and region between the second and third turning appearing different in the PDA dataset from the other two. The fact that the on-screen dataset looks more like the dataset where no info-marks were presented rather than the PDA dataset suggests that users with the on-screen delivery method navigated more like those who had no info-marks presented than those who had them delivered via the PDA.

Obviously the particular features seen here would be interpreted differently depending on what information was being displayed by each visualisation technique (elevation, colour and shading), but using the characteristics identified in the visualisation in this particular study assumptions could be made about the landscape from track logs collected from a different VE. For example, if we saw in the data a mountain range with a trench through it we would be able to determine that in that location there was a road that users followed. Flat areas in the data would correspond with flat areas in the environment and tall data peaks would generally correspond to peaks in the environment. We would also be able to identify key check points in the environment by a more intense colouring and shading introduced by high levels of track counts and congestion.

Although specific features were identified in terms of the elevation in the visualisation, these are all based on the IDW diagram derived from the clustering dataset. If we were to use a different dataset as the elevation aspect the characteristic features would have a different meaning. For example, if the track count data was used (Figure 70), then the display would be stepped owing to the discrete values used. Data-mountain ranges in this instance would indicate a common route through the environment where multiple users travelled, data-peaks would signify intersections of multiple tracks moving in different directions and data-valleys would be caused by regions where no or few tracks went.

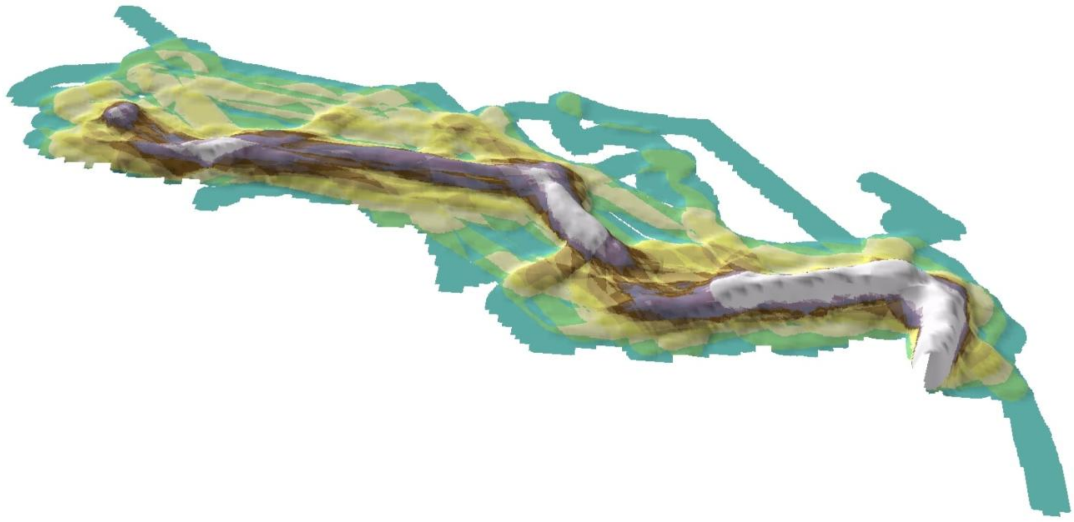


Figure 70: Track count as elevation

Similarly displaying congestion as the elevation changes what the different features imply (Figure 71). Owing to the nature of the analysis using percentages along the tracks we see a bobbled effect throughout the data which is not caused by changes in the tracks themselves but by the boundaries of the percentage regions (i.e. in between the 56 and 57% there would be a dip where the boundaries of both were less than the central kernel of each). We still see the familiar data-mountain ranges and a higher extent of data-peaks, but owing to the wider areas that the elevation data is generated from (there is no interpolation suddenly between high and low values) data-ridges become more difficult to generate.

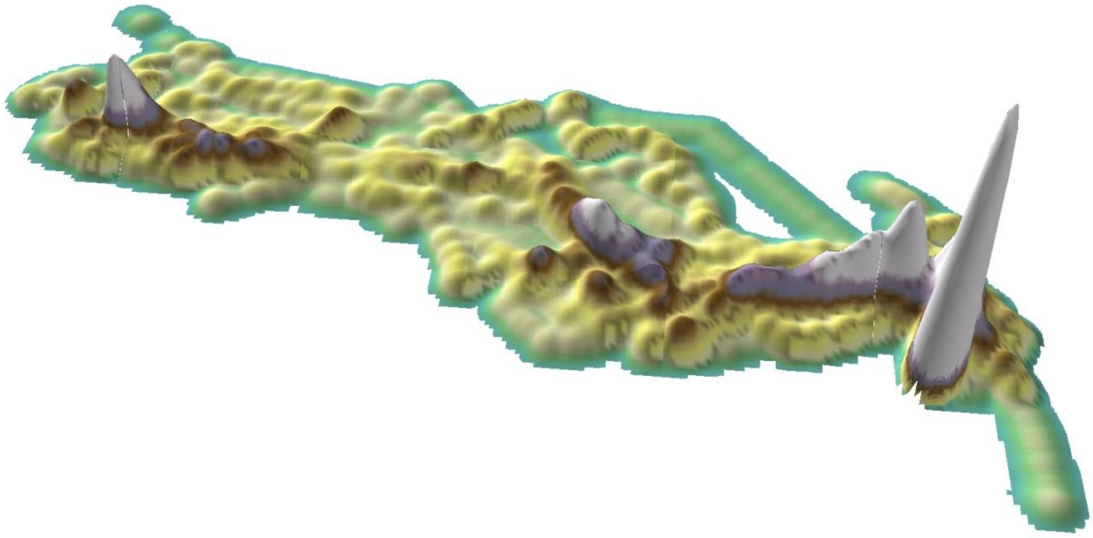


Figure 71: Congestion as elevation

Additional advancements are with regard to the visualisation of both the data and environment together. In this study it was not possible to display both at the same time owing to licensing restrictions with the Blueberry3D software used and so individual screenshots were compared instead. A future study could investigate the possible interaction techniques with the two visualisations at the same time i.e. whether they should be interacted with individually or movement in both simultaneously, what visual aids could be used (maps, compass etc.) and whether visual and HCI techniques have an impact on visual analysis (stereo vision, treadmill movement etc.).

In conclusion and in answer to the research question ***"****Can a multi-metric analysis be performed on track log data to derive information about trial performance and the environment itself?"* it is clear that by visualising the three pieces of data at the same time alongside the actual environment has been of great benefit in terms of identifying what features have an effect on the navigational performance. By combining the data into one visualisation it was possible to distinguish distinctive features in the data which signified either problem areas or useful features in the environment. In addition this investigation has shown that this visualisation method can be used to compare datasets and identify differences between them which would be missed by simple statistical analysis. Overall it can be seen that it is important to take into account multiple aspects of the track log data rather than a single numeric value. Although this numeric value can tell us whether there has been a significant difference between user's tracks it does little to tell us the reasoning behind it. Therefore it can be said that the multi-metric analysis performed here can tell us more about the user performance than the numerical analysis and individual analysis of each dataset derived. Obviously analysing the datasets side by side rather than as a single visualisation could lead to similar conclusions about the navigational performance, but it is expected that it would be more difficult to match areas of interest between the datasets, especially when the data is scaled.

8 Conclusions and Discussion

Throughout this thesis, several aspects relating to spatial task performance within virtual geographic environments have been investigated. This section provides an overall summary of findings, alongside a discussion identifying how these findings relate to previous research. Next an evaluative discussion of the thesis including the merits and short fallings of the study as a whole is presented, before concluding with proposed future studies and a general summary.

8.1 Summary of Findings

Within the research presented in this thesis, several findings have been made with respect to the implementation of tools to aid in spatial performance tasks, and methods for collecting and assessing the impact of such tools.

Firstly, overview maps were included into the VE as an aid to improve the estimation of observational long range distances. Owing to their use in studies such as Meilinger & Knauff (2008) whereby overview maps are used to construct spatial knowledge of an environment, it was believed that the inclusion of such overview maps would help to create a more accurate spatial representation of the environment. Statistical evaluation of distances provided showed that the inclusion of such overview maps did not alter the accuracy of such estimations and neither did the ability to zoom in on the overview map. This is counter-intuitive due to their ability to aid in the construction of spatial knowledge (Meilinger & Knauff, 2008), although this may be a result of the storage method of the mental representation (Meilinger & Knauff (2008) note that the route shown on a map may be stored as a series of directions). Unlike previous research where the accuracies of distance estimations were assessed (i.e. Clément et al. (2008) and Bodenheimer et al. (2007)), there was no bias towards underestimation of distances and there was an almost equal number of overestimates to underestimates. Analysis of qualitative information gained from in-trial feedback and post-trial interviews did suggest however that the users felt that when the map was provided their estimations were more accurate and the task easier. This discrepancy between results could be due to the test subjects making use of the maps but not being able to ascertain accurate information from them.

Also investigated was whether the inclusion of a map in one instance aided in estimations in a second test where the map had been removed. Owing to the use of VEs in training for navigational tasks (i.e. Farrell et al. (2003), Grant & Magee (1998)) it as expected that skills would be learnt in the first estimation which would then transfer to the second. Quantitative data indicated that there was no evidence of such an effect but the qualitative data indicated a difference in perception of ease. When the map was provided in the first estimation, the subsequent estimation when it was removed was found to be more difficult than when there was no map present in the first estimation. This finding may be due to the users relying too heavily on the overview map when it was presented and so when it was removed there were no skills gained from the first estimation that could be transferred. When no map was present in the first instance, people made use of the environment to estimate the distance and so information gained in the first estimation could be directly transferred. The initial belief was that the inclusion of the map in the first instances would allow for the transference of skills to the second, as is seen when moving between virtual and real world instances (i.e. Grant & Magee (1998) and Lathrop & Kaiser (2005)). This transference was not evident in this investigation and so it may be the case that either the tool was not useful in such a context, or that the method of transference is different between tasks than between environments.

Overall, from the distance estimation trials it was found that although the inclusion of the overview map (in either static or scalable form) did not alter the accuracy of distance estimations, it did appear to have an effect on the *perceived* accuracy and ease of task. In addition, when the map was removed users found subsequent estimations more difficult as they relied heavily on the map to determine the distance information. The predominant contribution of this research is the setting in which the estimations were performed in – a large scale rural environment. Most previous research has used either structured or small scale environments, with the distances that are estimated being relatively small. The tools and investigative methods implemented in the research presented in this thesis however are targeted at larger distances in a fairly unstructured environment, a scenario which appears to have been predominantly neglected. In addition, this study did not aim to understand why estimated distances are erroneous (as investigated in studies such as Witmer & Kline (1998) and Interrante et al. (2008)) but instead aimed at attempting to improve the estimation by providing a visual tool, an area that appears to be under-researched.

The second portion of the experimental area of the thesis investigated the effects of providing areal 'info-mark' features to users as a means of improving navigational performance. By recording track log information from the users as they moved through the environment in an attempt to follow a route shown on a paper map, several metrics were obtained which were used as a basis for performance analysis.

To investigate the effects of such a tool, statistical testing of the derived metrics was undertaken. This testing identified that the info-marks did little in the way of altering performance within the task, with the only statistically significant result being found within the overall average travel speed. This result was between the instances where no information was provided and when it was delivered via a PDA device. That result identified that when presented with the information via the PDA users travelled on average faster through the environment than when it was not provided at all. Qualitative analyses of results, on the other hand, indicate that when they were presented, the info-marks were used heavily with several users stating that they were the primary information used to determine their route through the environment. The difference found here between the qualitative and quantitative results indicate that indeed the info-marks were used to aid in the following of the route, but when they were not provided users could make use of other features within the environment to navigate. Similar to the studies by Counts & Smith (2007) and Field & O'Brien (2010), it appears that the inclusion of the geo-located information does alter the user interaction with the environment. The evidence from this investigation does however appear to suggest that this is more of a perceptual change rather than one that actually alters the performance within tasks.

It was found however that there was a correlation between the level of perceived skill and performance, and the area metric used in the analysis. This correlation showed that if users rated their performance and skills as being higher, the area between the intended and actual route decreased. This therefore could be a useful metric in identifying performance in navigational tasks where users are asked to trace a route. Again a primary contribution to the research field from this study is the implementation and investigation of tools developed for use in an unstructured and large VE rather than the more often used indoor environments or those where the direction of movement is limited. Bidwell & Lueg (2004) state that natural environments may be qualitatively different from structured man made ones, and so the research into natural environments is important.

Another aspect investigated as part of the navigation task was to identify what features within the rural environment could map to the feature types identified in the UIT. As the types can be used as a means of classifying features within urban environments and then these used to improve navigation within that environment, identifying similar features within rural areas may be of benefit for improving navigation in such rural regions. As Vinson (1999) stated that all five feature types should be included in an environment, it is important in the context of rural environments to identify what features actually fall into these categories. It was identified that within the environment used, 'classic' features that fall into the classifications are seldom present and so additional rural features were identified. These included the identification of rivers and valleys as paths, crop plantations as districts and steep slopes as edges. A map was generated using the classifications generated that highlighted such features within the environment which may be of benefit when navigation tasks are performed. As implemented by Omer et al. (2005), the classifications of features were identified in terms of 'hardness' (how many people identified the feature as a particular type) which meant that the most 'hard' features were used to generate a map of the features rather than all features identified.

The final research aspect of the thesis was the investigation of a visualisation technique for portraying multiple pieces of information about the track logs acquired in the navigation task. Unlike numerous other studies such as those by Chittaro et al. (2006) and Drachen & Canossa (2009a), the visualisation developed in this study presented multiple pieces of information in a single visualisation using data from a larger number of individuals. By producing a 3D surface with a raster draped over it, it was possible to investigate three aspects relating to the track logs at the same time: clustering, congestion and track count. With this technique it was possible to identify key areas within the data as well as to draw conclusions about general performance.

Several data features were identified which signified particular performance traits including the 'data-plateaux' which identified constant travel speed, 'data-mountain ranges' which showed areas of a high number of track counts and slow movement, and 'data-peaks' which identified areas where users paused for a time. Each of these data features can be used to highlight areas within the environment where users found movement easy or orientation difficult. In addition, when using the visualisation technique on the three datasets (no info-mark, on-screen delivery, and PDA delivery), differences could be seen between the overall 'look' of the data. In particular, the PDA dataset appeared more dispersed and flatter than the other two which implied that a different navigation technique was used in those tracks. The overall benefit of the visualisation section was the determination that the method used was of assistance when assessing overall navigation performance. In addition, particular data features found mapped to real world environmental features which suggests that the data visualisation can actually be used to determine information about the environment that the data was collected from. A number of these mapped real world features fit with the UIT classifications discussed, and so links can be made with the data visualisation and how people perceive the environment. For example, data-peaks tend to be found at node type features which identify that such data-peaks occur at key navigation checkpoints. This visualisation showed that the visualisation of track log information does not need to be constrained to the presentation of singular metrics for large datasets, or the use of smaller datasets when multiple metrics are displayed at once (i.e. Hoobler et al. (2004)).

8.2 Discussion

As mentioned, this research has identified several key findings. These are predominantly methodological in nature in that they emphasise the need for both qualitative and quantitative data collection and analysis. Throughout the study contradictions were found between users' perceived performance (identified through qualitative data) and their actual performance (quantitative data). If only one of these aspects was investigated then it is likely that inaccurate conclusions would be made.

Several other findings relating to analysing information and improving navigation have been presented such as the use of the multi-metric track log visualisation and the identification of different rural features that match the UIT feature types. Such findings can be of great benefit for analysing and understanding navigational performance and for aiding in the development of environments which are more navigable and easier to understand.

In the context of this study, it was felt that the user focussed activities conducted were the most suitable approach to identify the effects of the tools implemented. With respect to estimating distances, although it was identified that verbal reporting may not be the most accurate (Interrant et al. 2006), other methods such as blind walking were simply not practical over the distances implemented. In the navigation task users were asked to follow a route depicted on a map which meant that map reading skills played an important role even though the use of a map was not the aspect being investigated. Obviously the map (or similar method) was needed to depict the location of features the info-marks referred to and so the use of the same map to show the intended route was logical. Other methods could be used however such as the user being taken through the environment on a route and then asked to repeat it, but it was felt that this would be more of a memory exercise. Also it was felt that the depiction of a route on a map was more realistic in terms of what methods would be used in a real world scenario.

Although several key findings have emerged from the research, these are subject to a number of limitations. Firstly, due to time constraints and access to potential test subjects the sample sizes used for the investigations were relatively small. This resulted in a limitation in the statistical analysis that could be performed as well as an increased impact of outliers on any results. The analysis of collected data used statistical tests and analysis methods that were tailored for such sample sizes and outliers (i.e. t-tests and Spearmans Rank) but the small sample sizes may account for a number of non-significant results.

Secondly, test subjects were drawn from a particular group of people (staff and students within a geography department) which may introduce bias if the sample was used as a representation of the general population. Therefore any results obtained from this thesis should be considered as a reflection of people who are used to handling geographic data and not the general population. Indeed it may well be the case that a number of the findings do transfer to the general population but this cannot be assumed without further investigation using test subjects from numerous backgrounds and areas.

Finally, although the test subjects were drawn from a population of geographically minded individuals, few controls were added to identify different skill levels and preferences between the test subjects. As noted in the literature review, different people have varied preferences on learning style (Hawk & Shah, 2007), and thus these preferences could have an impact on the effectiveness of tools developed. Voyer et al. (1995) provide a list of studies that make use of such spatial ability assessment tests (13 different tests were identified in the list) with regards to assessing performance in a number of tasks including spatial perception, mental rotation and spatial visualisation. It would be beneficial to implement one or more of these tests as a means of quantifying spatial performance of individuals but there was neither the time nor resources available to perform such analysis. An additional time consideration with respect to the use of such tests is the identification of an appropriate test to perform which would require extensive research. The issue relating to a lack of such controls is that there may be sub-groups within the sample that perform better or worse in the tasks, or that make use of the tools provided differently. Without a thorough investigation to identify (and control for) such sub groups the collective differences introduced by these may affect the overall results obtained. Therefore further investigation to identify such groups should be undertaken for any future analysis.

8.3 Possible Modifications

Due to the presence of limitations and general findings from the studies conducted, several improvements and modifications can be identified which would improve the investigation if it was to be repeated. Such improvements predominantly revolve around the sampling methods employed and the design of tools implemented.

Firstly, as identified earlier, limitations are present in the study due to the limited sample size and the fact that the test subjects were drawn from a specific group of people. To improve the investigation, a larger sample size should be implemented so that the results obtained are not influenced as strongly by outliers and any patterns would emerge. The predominant reasoning behind the limited sample was time constraints in that the number of trials that could be run was limited by the overall time allocated for data collection, collation and analysis. In addition, temporal issues were introduced due to the time of year that experiments were conducted where students may be on holiday or too busy with study commitments to take part in the investigation. However, it was felt that the use of both qualitative and quantitative data collection and analysis was important. This resulted in more time being required for each trial and the subsequent analysis of data, and so the collection of more data would require extending the overall time allocated to the data collection portion of the study.

Another aspect that possibly limits the overall transference of the results found to the general population is the decision not to collect in depth information regarding spatial literacy and spatial ability of test subjects. Pilot studies were conducted to assess aspects including overall time required for a trial, and due to time constraints and proposed sample size it was decided that such in depth investigation would not be feasible in the project timeframe. Although this may limit the applicability of the findings to the general population, by selecting participants from the intended end-user group (geography students and staff) any tools developed would be relevant to those people. To make the results transferrable, more participants would be required from a wider sample pool, as well as consideration into performing the in-depth investigation into spatial literacy/ability. In future investigations this would be more feasible as the testing environment has already been developed in this study.

A further aspect identified from the qualitative feedback of the experiments was that the tools developed may not be the optimal method of providing information to the user with regards to the spatial tasks undertaken. In the case of this study, only one design and method of delivery (map and info-mark) was investigated whereas in future studies a comparison between methods would be beneficial. For example, in the instance of the info-marks an audio delivery method may be more useful than the visual text based representation. In general, aspects such as delivery methods, user interaction and general design may have an impact on the usefulness of the tools implemented and so these are aspects that warrant further investigation. Although considerations were made with respect to these aspects, more in-depth investigation into the effects of different implementations on such exercises may have been of benefit at the design stage. However, an investigation on the impact of such design considerations (e.g. differences between visual/audio delivery methods, size of the overview map) would be a large research topic in itself and so is far outside the scope of this investigation. That said, it is a foundation problem in that if the wrong design of tools are implemented then any results obtained from investigations using those tools will not be optimal. Therefore as an extension to the research conducted in this thesis, an investigation into the differences in performance introduced by the altering of designs and implementation would be of benefit. If a larger timeframe was available for the research presented here, these considerations would have been directly implemented in the investigations conducted.

Overall, these proposed changes highlight that it is important to consider design aspects (in both tool and sampling contexts) before undertaking research in which results may be affected by user preference and experience. A predominant contributing factor to the reasoning why such considerations were not implemented fully in this thesis is that of time constraints in that there simply was not enough time to investigate the chosen topics and to fully investigate and implement the aforementioned considerations. If such considerations were to be undertaken then the research questions posed at the start of the study would need constricting to allow for the extra time required for the additional investigative aspects.

8.4 Further Study

As discussed in the previous chapters although the research questions posed have been answered, results from the studies have posed several new questions. Although outside the scope of this study, the questions derived could provide more in depth understanding into several aspects regarding virtual reality, distance estimation, navigation and data visualisation.

8.4.1 What Effects do Different Landscapes Have on Subsequent Distance Estimations?

Although not statistically significant, results from the transfer of knowledge task undertaken during the distance estimation experiment appeared to indicate that there may be some link between distance estimations in different environments and the features present within them. In this study, data suggested that users who conducted their initial estimation in one particular location performed less accurately in a subsequent estimation than other users at different locations. Therefore, what in particular was it about that first location that made the subsequent estimation less accurate (if indeed it was) than those provided by users from a different location?

8.4.2 Does the Presentation of Speed andDistance Travelled Information Improve Navigational Performance?

From feedback provided in post-trial interviews during the navigation experiment it was made clear that users were finding it difficult to identify how far they had travelled within the environment. Several responses were given as to how this could be rectified, with a general theme emerging that revolved around the display of either information regarding distance travelled or a car like speedometer visualisation. If either of these tools were implemented in a subsequent environment, would user’s navigation performance increase as a result of knowing how fast they are travelling or direct information regarding distances travelled?

8.4.3 Do Different Locomotion Methods Within a VE Alter the Features Classified as Urban Image Theory Elements?

When identifying features within the rural environment that could be classified into the UIT classifications, some responses indicated that the method of how people travelled within the environment and their interaction with the features could have an effect on their classifications. This is synonymous to the ‘is a road a path or an edge’ classification in an urban setting where a driver would see the road as a path yet a pedestrian may see it as an edge. If interaction with the VE was altered (by using a car ‘cockpit’ or treadmill) then would this alter how people classified features? Also, would placing restrictions on movement depending on features (i.e. not crossing rivers, different travel speeds based on inclination etc.) have an effect on the classifications?

8.4.4 Can the Multi-Metric 3D Visualisation be Used on Track Logs from an Unknown Landscape in Order to Derive Information About the Environment Itself?

During the visual analysis of the track log data, several data features were found to correspond with actual features within the environment. Such links included data-peaks at peaks within the environment, data plateaux corresponding with large flat areas in the environment, and data trenches within ridges occurring along roads. Therefore, if track logs were obtained from a similar task in a different environment, could the analyst derive a general ‘snapshot’ of the environment itself by linking data features with features that could appear in the same location?

8.5 General Summary

In this study several research questions have been posed and answered via the qualitative and quantitative analysis of user performance within tasks. Although some results indicated that methods implemented had little effect on user performance, the methods used show that it is important to address such problems using different analysis techniques. Without performing the collection and analysis of both qualitative and quantitative data, several aspects would have been missed including the fact that users felt that a task was easier and that they were more accurate even though this was not the case. In reality, the actual lack of evidence to reject the null hypothesis for many of the questions is interesting. For example it should be expected that when presented with discrete information indicating distances then users should be more accurate at estimating such distances. Also when presented with information directly linking to features present on a map also identifying an intended route, it was believed that users should be able to follow the route more accurately as they have extra information. In both these cases there was not enough evidence to suggest this was true, so why was there not? Is the lack of benefit to do with user preferences, delivery methods, or simply that the tools used were not suitable? Indeed it is possible that the small sample sizes used are the reason behind the non-significant results, but this cannot be said for sure without further investigation.

Several positive results have been found from this study however, such as the benefits of using the multi-metric visualisation method to analyse track logs, the ability of the methods used to collect data and to provide information to the users, and the identification of features within the rural environment that can take the place of Urban Image elements. By visualising the track log data in the method implemented it was possible to see that although statistically there was very little difference in task performance between the groups, we could see that in the PDA group there appeared to be different navigational techniques being used. Whereas the on-screen and no info-mark datasets looked fairly similar, the PDA dataset was visually different appearing flatter and more dispersed. Without the visual analysis, this difference would have been overlooked. Finally, the features identified that matched Urban Image classifications could be of great benefit for the future development of rural virtual and real world environments. Although obviously it is difficult to manipulate the terrain in a real environment (you cannot just create a valley) by knowing that the features identified form the basis for a rural image, these elements could be highlighted on a map to aid in navigation.

Another aspect of interest identified was that although the presentation of additional information had little effect on actual performance, the perceived performance by the user was different. In most cases, users felt that the task was easier and that they were more accurate when the additional information was provided (overview maps of info-marks via the PDA) even though this was not the case. Therefore if such methods are implemented in tasks where the users' perceived performance and accuracy was important (i.e. the determination of walking distance at dusk discussed earlier) care would be needed to ensure that the users did not become too overconfident in their erroneous measurements.

Overall the primary findings from this study are:

- Providing overview maps does not improve estimations of observational distances;

- Users felt more accurate and that the distance estimation task was easier when presented with an overview map;
- The provision of info-marks had little effect on route tracing performance, with the only statistically significant result being in the average travel velocity;
- An increased level of accuracy and ease was perceived by users when info-marks were presented via a PDA;
- Multiple features within a rural environment can fit with the elements of the UIT;
- The use of a 3D multi-metric visualisation provided greater insight into navigational behaviour and the effects of features within the environment itself;
- It is important to use both qualitative and quantitative data collection and analysis methods to identify user feelings as well as to determine task performance.

In conclusion, this study has highlighted the importance of using both qualitative and quantitative data collection and analysis methods when investigating performance of cognitive tasks within VEs. It has also presented several novel visualisation and interaction methods for use in virtual environments, as well as techniques for visually analysing track log data.

9 Appendices

Appendix 1: Replicated Friedman Test Code 291

Appendix 2: Info-mark Text 294

Appendix 3: Track log Raw Data 301

9.1 Appendix 1: Replicated Friedman Test Code

```
location <- factor(c("Location 1","Location 2","Location
3")[c(1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3,3)],ordered=TRUE)
map.type <- factor(c("No Map","Static Map","Scalable
Map")[c(1,1,1,2,2,2,3,3,3,1,1,1,2,2,2,3,3,3,1,1,1,2,2,2,3,3,3)],ordered=TRUE)

# This function computes a replicated version of the Friedman stat - based on his paper
rep.friedman.stat <- function(score,treatment,block) {
      treatments <- unique(treatment)
      blocks <- unique(block)
      rank.score <- integer(length(score))
      for (b.level in blocks) {
            items <- which(block == b.level)
            rank.score[items] <- rank(score[items])}
      rhj <- tapply(rank.score,list(treatment),mean)
      rha <- mean(rank.score)
```

```
      sst <- length(treatments)*sum((rhj - rha)^2)
      sse <- sum((rank.score - rha)^2) / (length(treatments) * (length(blocks)-1))
      return(sst/sse)}

monte.carlo.friedman.stat <- function(treatment,block,nsims) {
      result <- numeric(nsims)
      treatments <- unique(treatment)
      blocks <- unique(block)
      rank.score <- numeric(length(treatment))
      for (i in 1:nsims) {
            for (b.level in blocks) {
                  items <- which(block == b.level)
                  rank.score[items] <- sample(length(items))}
            rhj <- tapply(rank.score,list(treatment),mean)
            rha <- mean(rank.score)
            sst <- length(treatments)*sum((rhj - rha)^2)
```

```
        sse <- sum((rank.score - rha)^2)/(length(treatments)*(length(blocks)-1))
        result[i] <- sst/sse}
    return(result)}

# Perform the test
monte.carlo.friedman <- monte.carlo.friedman.stat(map.type,location,10000)
actual.friedman <- rep.friedman.stat(score,map.type,location)

# and work out the proportion of times the actual value exceeds the monte carlo simulations -
this is the experimental p-value
sum(actual.friedman > monte.carlo.friedman)/10000
```

9.2 Appendix 2: Info-mark Text

Name	Information Provided	Location Ref.
Iglesia de Santa Maria	The Iglesia de Santa Maria is a church located in Sorbas. Built in the sixteenth century, this church sits in the main square of Sorbas and was built on the site of an ancient mosque.	0
Cantona Peak	Cantona Peak is one of the highest points in the area, reaching over 755 metres. Located at the top of the peak is a cap stone marking the height of the location.	1
Urrá	Urrá is an estate in the Sorbas basin offering accommodation and field trips. Located near the river and the local caves, it is in a prime location for geologists and students looking at investigating the area.	2
Los Perales	Although located very close to the Autovía del Mediterráneo, Los Perales has no direct access and residents must travel all the way to the junction with the Old Sorbas Road to gain access to it.	3
Peñas Negras	Peñas Negras is a village located on the Old Sorbas Road close to the Autovía del Mediterráneo. This area is also a good lookout point to view the different sediments that fill up the Sorbas Basin	4
Mizala	Mizala is a village located on the flat plateau region between two peak ranges within the area. The surrounding terrain is predominantly agricultural, providing a patchwork terrain of cultivated regions.	5

Los Molinos	Los Molinos is an abandoned village in the driest area of Spain, although it is close to the only river valley in the country that runs all year round. Los Molinos literally translates to 'The Mills'	6
Sorbas	Sorbas is the largest town in the area, sitting above a meander in the Río Aguas. The name 'Sorbas' is Arabic in origin, literally meaning 'pot of sand'. No. Of inhabitants 2,854	7
Huelí	Huelí is located deep in the hilly central area of the Sorbas region, with the main entrance road passing two gypsum quarries. The surrounding terrain is dominated by a layered appearance, making the slopes of the hills appear more like large staircases.	8
Larache	Larache is a small collection of buildings located just outside of Sorbas. Situated close to an industrial ground, the buildings also back on to cultured vegetation patches.	9
Los Contreras	Los Contreras is a small estate located in the vegetated hills of the region. The only access road to the estate comes from Sorbas, cutting up the hillside.	10
Los Molinos del Río Aguas	Situated between Old Sorbas Road and the Río Aguas valley, Los Molinos del Río Aguas translates to 'The Mills of Río Aguas'. Water springs created from the Río Aguas have formed an oasis like appearance to the area, with the village being surrounded by lush greenery in the desert environment.	11

El Tesoro	El Tesoro translates to ‘The Treasure’ and is a collection of buildings lying in ruin next to a small valley.	12
El Tieso	The translation of El Tieso is ‘The Stiff’ and is a small collection of structures located near the Río Aguas valley.	13
El Mayordomo	El Mayordomo is a small village located in one of the hilly regions of the area. Situated above one of the Río Aguas valleys, there are many plantations surrounding the village.	14
Quijiliana	Located near a meander in the Río Aguas, Quijiliana is a village surrounded by plantations and vegetation. On the other side of the river valley the land is far less cultured.	15
Pocatorta	Pocatorta is an estate situated on a peninsular in the hills where two riverbeds join.	16
Góchar	Another of the smaller villages in the area nestling close to the Aguas valley is Góchar. North of the valley, the surrounding terrain is somewhat more hilly with vegetation being more prevalent and greener.	17
Barrio de las Alfarerías – Pottery District	The barrio de las Alfarerías, or Pottery District, is the predominant home of many of Sorbas’ pottery workshops.	18

Gypsum Karst	Gypsum karst is a type of landscape formed by the dissolution of a karst bedrock layer. Through time (taking thousands of years) the dissolving of the subterranean rock results in the formation of caves and depressions in the surface. Karstification of gypsum does not happen very often in nature, with limestone karstification forming the greater part.	19
Old Sorbas Road/Freeway Junstion	In this area there are two main roads – the Autovía del Mediterráneo and the Old Sorbas Road, the Autovía del Mediterráneo is a motorway type road running all the way from Algeciras near Gibralta to Barcelona, some 1300km. Old Sorbas Road is a winding lesser road running through the Sorbas Basin, connecting the main carriageway to Sorbas.	20
Road to Campico	Located near the main highway 'Autovía del Mediterráneo', in this area there is a junction between the A-1102, or Old Sorbas Road as it is known and a road leading to the village of Campico.	21
Shale	Close to the Río Aguas in this area, the ground surface turns purple from the shale present. Shale is a sedimentary rock, created from mud and clay deposited in layers, forming the layered structure apparent in the rock deposits.	22

Barranco del Infierno and Cueva del Yeso	The Barranco del Infierno (Hells Gorge) and Cueva del Yeso (Yeso Cave) are gypsum karst features created by fluvial erosion and karst dissolution. The cavern was formed at the meeting point of two karst barrancos where at some point in time, a sinkhole was established allowing the capture of water which flowed down the barrancos. This meant that not all stretches of these fluvial courses continued to deepen, resulting in some parts of the barranco remaining higher than others, and the water flowing underground formed the cave system.	23
Fossil beaches of Sorbas	Located near the town of Sorbas, these fossil beaches are mud and sand sediments deposited around 5.4 million years ago. Formed on a system of barrier islands, these beaches separated an enclosed shallow lagoon of mud and silts and a bay formed by the ancient Mediterranean Sea.	24
Lapiés	In this area, lapiés can be found which are sharp crested grooves and furrows created by the dissolution of gypsum.	25
Gypsum Quarry	A predominant rock type found within the Sorbas region is Gypsum. It is so abundant in places that quarries have been built to extract it from the ground for use in applications such as making plaster for buildings. Within this area, two of the large quarries can be found.	26

Abandoned Valley	In this area the canyon formed by the Río Aguas can be clearly seen. The sides of the valley are particularly unstable here with slipped masses of Azagador Member limestone present, as well as terraces. Prior to the building of the Autovía del Mediterráneo, sections revealing around 10 metres of coarse imbricated fluvial gravels set into the Tortonian Marls were present.	27
Terrace	The river terraces that can be observed in the valleys in the area are formed when the elevation of the rivers and flood plains decreases suddenly. This results in a layering effect with the new river level cutting into the previous bed.	28
Road to Los Canstaños	A junction between two of the main roads in the area – Old Sorbas Road (A-1102) and the N-340a can be found close by. Heading north on the N-340a will eventually lead you to Los Canstaños, passing by quarries for gypsum.	29
Road to Huelí	Within this area, a junction on the Old Sorbas Road can be found that runs to the village of Huelí, passing two gypsum quarries.	30
Road to Uleila Del Campo	A junction between the Calle de Alcalá and the AL-4101 allows for the northward travel to Uleiel del Campo, a town similar in area to Sorbas with a population of 1,015 inhabitants in 2009.	31
Road to El Mayordomo	Located within this area is the junction between the AL-4101 road to Uleila del Campo and the smaller road leading to the village of El Mayordomo.	32

Dolinas	Dolinas are circular depressions found on the ground which are created from wells or sinkholes which connect them to subterranean caverns of gypsum karst. There are several ways in which these Dolinas can form – cave-ins of a gypsum level, erosion of a marl level, or the dissolution of a gypsum level.	33
Túmulos	Túmulos, which can be found in this area, are mound like features exclusive to the Sorbas region. They form from the doming of the surface layer of gypsum, generated through an increase in the volume of water absorbed by the gypsum crystals.	34

9.3 Appendix 3: Track Log Data

Trial	Gender	Age	Map skill	Navigation skill	Ease	Accuracy		Length	Length Error	Area	Time	Speed	Distance from end	Combined
1	Female	15 - 20	7.5	7.5	6	5		6743.39	403.31	1543262	1028	6.56	459.97	-0.40
2	Male	15 - 20	4	4.5	5	3.5		6716.45	376.38	1616616	936	7.18	465.10	-0.32
3	Female	21 - 30	7	7	N/A	4.5		6628.64	288.56	1131223	931	7.12	59.06	-1.71
4	Male	21 - 30	8	8	7	5.5		7338.13	998.05	455732.3	958	7.66	478.02	-1.22
5	Male	21 - 30	7	8	8	6		7381.68	1041.60	596543.2	903	8.17	318.90	-1.28
6	Female	15 - 20	6	7	8	7		6395.48	55.40	899874	713	8.97	186.73	-1.98
7	Female	31 - 40	6	6	6.5	3.5		6767.33	427.25	1083530	1974	3.43	86.25	-1.60
8	Female	31 - 40	7	6	6	7		9605.46	3265.39	1659607	804	11.95	354.83	2.02
9	Female	15 - 20	6.5	6.5	7	6		6552.55	212.47	966695	960	6.83	147.02	-1.83
10	Female	51 - 60	9	2.5	5	7		6985.14	645.06	1053532	1205	5.80	139.42	-1.36
11	Male	21 - 30	8.5	8.5	4.5	6.5		8428.11	2088.04	374421.4	1702	4.95	84.34	-1.06

12	Male	15 - 20	7	8	4.5	7		7646.13	1306.05	1733854	935	8.18	499.21	0.68
13	Female	15 - 20	8	7	3	7		6857.13	517.06	424301.8	1297	5.90	61.08	-2.39
14	Female	31 - 40	8.5	8	4	6.5		9208.72	2868.64	967545.1	2013	4.57	69.06	0.32
15	Female	31 - 40	7	6.5	8	7.5		5982.22	357.85	740868.8	1054	5.68	503.61	-1.37
16	Female	21 - 30	7	7	5	3		7053.10	713.03	834671.8	1271	5.55	376.65	-1.16
17	Male	31 - 40	8	8	6	4		8188.37	1848.30	1074493	1143	7.16	107.67	-0.35
18	Female	21 - 30	9	9	7	8		6094.03	246.05	589523	1282	4.75	273.68	-2.05
19	Male	41 - 50	8	8	2	2		10759.13	4419.06	2381996	1745	6.17	59.47	3.41
20	Male	31 - 40	8.75	8	6	7		7060.32	720.25	370691.8	1483	4.76	42.32	-2.31
21	Male	21 - 30	8.5	7	5	6		7353.11	1013.04	2680715	832	8.84	1255.16	2.92
22	Male	31 - 40	9	5	2	6		6567.01	226.93	391738.5	1465	4.48	29.12	-2.74
23	Female	15 - 20	3	3	5	6		10068.81	3728.74	2563801	1884	5.34	72.48	3.06
24	Male	31 - 40	7	8	7	8		7045.77	705.69	600414.3	1002	7.03	60.61	-2.01
25	Female	21 - 30	5	5.5	3	3		8487.12	2147.04	997144.6	1325	6.41	991.57	1.34
26	Female	31 - 40	7	4	4	5		7725.96	1385.88	1820169	1260	6.13	65.23	0.11

27	Male	21 - 30	8	6	5	7		7532.88	1192.80	725121.8	1043	7.22	221.24	-1.16
28	Female	21 - 30	5	3	3	4		3661.35	2678.72	3624558	563	6.50	2985.53	8.52
29	Female	21 - 30	7	8	8	6.5		6740.14	400.07	641473	1031	6.54	386.87	-1.66
30	Female	31 - 40	5	4	4	6		15116.13	8776.06	2295517	2407	6.28	351.77	7.55

References

Abu-Obeid, N. 1998, "Abstract and scenographic imagery: The effect of enviromental form on wayfinding", *Journal of Environmental Psychology*, 18(2), pp. 159-173.

Adrienko, N. & Adrienko, G. 2011, "Spatial generalization and aggregation of massive movement data", *IEEE Transactions on Visualization and Computer Graphics*, 17(2), pp. 205-219.

Ahmed, F., Cohen, J.D., Binder, K.S. & Fennema, C.L. 2010, "Influence of tactile feedback and presence on egocentric distance perception in virtual environments", *IEEE Virtual Reality Conference*, VR ’10, pp. 195-202.

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I. & Angel, S. 1977, *A Pattern Language*, Oxford University Press, New York.

Alexander, T., Conradi, J. & Winkelholz, C. 2003, "Depth perception and visual after-effects at stereoscopic workbench displays", *Proceedings of the IEEE Virtual Reality*, VR ‘03, pp. 269.

Allison, R.S., Gillam, B.J. & Vecellio, E. 2009, "Binocular depth discrimination and estimation beyond interaction space", *Journal of Vision*, 9(1), pp. 1-14

Aoki, H., Oman, C.M., Buckland, D.A. & Natapoff, A. 2008, "Desktop-VR system for preflight 3D navigation training", *Acta Astronautica*, 63(7-10), pp. 841-847.

Appleton, K. & Lovett, A. 2005, "GIS-based visualisation of development proposals: Reactions from planning and related professionals", *Computers, Environment and Urban Systems*, 29(3), pp. 321-339.

Appleton, K. & Lovett, A. 2003, "GIS-based visualisation of rural landscapes: Defining 'sufficient' realism for environmental decision-making", *Landscape and Urban Planning*, 65(3), pp. 117-131.

Appleton, K., Lovett, A., Sünnenberg, G. & Dockerty, T. 2002, "Rural landscape visualisation from GIS databases: A comparison of approaches, options and problems", *Computers, Environment and Urban Systems*, 26(2-3), pp. 141-162.

Ardito, C., Costabile, M.F., De Angeli, A. & Pittarello, F. 2007, "Navigation help in 3D worlds: Some empirical evidences on use of sound", *Multimedia Tools and Applications*, 33(2), pp. 201-216.

Bertamini, M., Yang, T.L. & Proffitt, D.R. 1998, "Relative size perception at a distance is best at eye level", *Perception and Psychophysics*, 60(4), pp. 673-682.

Beyer, H. n.d., , *Hawth's Analysis Tools for ArcGIS*. [online] Available at: http://www.spatialecology.com/htools/index.php [Accessed 12th May, 2010]

Bidwell, N.J. & Lueg, C.P.. (2004). “Creating a framework for situated way-finding research”. In: Masoodian, M., Jones, S. &Rogers, B. *Computer Human Interaction*. Berlin: Springer Berlin/Heidelberg. 40-49.

Biggs, M., Fischer, U. & Nitsche, M. 2008, "Supporting wayfinding through patterns within procedurally generated virtual environments", *Proceedings of Sandbox 2008:- An CAN SIGGRAPH Videogame Symposium*, Sandbox ‘08 , pp. 123-128.

Bodenheimer, B., Meng, J., Wu, H., Narasimham, G., Rump, B., McNamara, T.P., Carr, T.H. & Rieser, J.J. 2007, "Distance estimation in virtual and real environments using bisection", *Proceedings of the 4th symposium on Applied perception in graphics and visualization*, APGV '07, pp. 35-40.

Börner, K. & Penumarthy, S. 2003, "Social diffusion patterns in three-dimensional virtual worlds", *Information Visualization Journal*, 2(3), pp. 182-198.

Bosco, A., Picucci, L., CaffÃ², A.O., Lancioni, G.E. & Gyselinck, V. 2008, "Assessing human reorientation ability inside virtual reality environments: The effects of retention interval and landmark characteristics", *Cognitive Processing*, 9(4), pp. 299-309.

Brooks, B.M. & Rose, F.D. 2003, "The use of virtual reality in memory rehabilitation: Current findings and future directions", *NeuroRehabilitation*, 18(2), pp. 147-157.

Brunsdon, C. 2009, "Replicated Friedman Statistic Algorithm" [Email] (Personal communication)

Brunsdon, C. 2012, "Friedman code equation" [Email] (Personal communication, 19th October, 2012)

Brunsdon, R., Nickels, L. & Coltheart, M. 2007, "Topographical disorientation: Towards an integrated framework for assessment", *Neuropsychological Rehabilitation*, 17(1), pp. 34-52.

Burigat, S. & Chittaro, L. 2007, "Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids", *International Journal of Human Computer Studies*, 65(11), pp. 945-958.

Caduff, D. & Timpf, S. 2008, "On the assessment of landmark salience for human navigation", *Cognitive Processing,* 9(4), pp. 249-267.

Chittaro, L., Ranon, R. & Leronutti, L. 2006, "VU-Flow: A visualization tool for analyzing navigation in virtual environments", *IEEE Transactions on Visualization and Computer Graphics*, 12(6), pp. 1475-1485.

Chittaro, L. & Venkataraman, S. 2006, "Navigation aids for multi-floor virtual buildings: A comparative evaluation of two approaches", *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST ‘06, pp. 227-235.

Claramunt, C. & Winter, S. 2007, "Structural salience of elements of the city", *Environment and Planning B: Planning and Design*, 34(6), pp. 1030-1050.

Clegg, F. 1982, *Simple Statistics: A course book for the social sciences*, Cambridge University Press, Cambridge, England.

Clément, G., Lathan, C. & Lockerd, A. 2008, "Perception of depth in microgravity during parabolic flight", *Acta Astronautica*, 63(7-10), pp. 828-832.

Counts, S. & Smith, M. 2007, "Where were we: Communities for sharing space-time trails", *Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems*, GIS ‘07, pp. 67-74.

Crompton, A. & Brown, F. 2006, "Distance estimation in a small-scale environment", *Environment and Behavior*, 38(5), pp. 656-666.

Cutting, J.E. & Vishton, P.M. 1995, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth" in *Perception of Space and Motion*, eds. W. Epstein & S. Rogers, Academic Press, San Diego, CA, pp. 69-117.

Darken, R.P. 1996, "Navigating Large Virtual Spaces", *International Journal of Human-Computer Interaction*, 8(1). pp. 49-72

Darken, R.P., Allard, T. & Achille, L.B. 1998, "Spatial Orientation and Wayfinding in Large-Scale Virtual Spaces: An Introduction", *Presence*, 7(2), pp. 101-107.

Darken, R.P., & Peterson, B. (2001). Spatial Orientation, Wayfinding, and Representation. *Handbook of Virtual Environment Technology*. Stanney, K. Ed., Lawrence Erlbaum Assoc.: New Jersey. pp. 493-518.

Darken, R.P. & Sibert, J.L. 1996, "Wayfinding strategies and behaviors in large virtual worlds", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '96, pp. 142-149.

Diamantopoulos, A. & Schlegelmilch, B.B. 1997, *Taking the Fear out of Data Analysis*, The Dryden Press, London.

Dockerty, T., Lovett, A., Sünnenberg, G., Appleton, K. & Parry, M. 2005, "Visualising the potential impacts of climate change on rural landscapes", *Computers, Environment and Urban Systems*, 29(3), pp. 297-320.

Dodiya, J. & Alexandrov, V.N. 2008, "Use of auditory cues for wayfinding assistance in virtual environment - Music aids route decision", *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, VRST '08, pp. 171.

Dolman, P.M., Lovett, A., O'Riordan, T. & Cobb, D. 2001, "Designing whole landscapes", *Landscape Research*, 26(4), pp. 305-335.

Drachen, A. & Canossa, A. 2009a, "Analyzing spatial user behavior in computer games using geographic information systems", *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*, MindTrek ‘09, pp. 182-189.

Drachen, A. & Canossa, A. 2009b, "Analyzing user behavior via gameplay metrics", *Proceedings of the 2009 Conference on Future Play on @ GDC Canada*, Future Play ‘09, pp. 19-20.

Elias, B. 2003, "Determination of landmarks and reliability criteria for landmarks", *5th Workshop on Progress in Automated Map Generalization*, Paris, 28-30 April 2003

Elvins, T.T., Nadeau, D.R. & Kirsh, D. 1997, "Worldlets - 3D thumbnails for wayfinding in virtual environments", *Proceedings of the 10th annual ACM symposium on User interface software and technology*, UIST ‘97, pp. 21-30

Farrell, M.J., Arnold, P., Pettifer, S., Adams, J., Graham, T. & MacManamon, M. 2003, "Transfer of Route Learning from Virtual to Real Environments", *Journal of Experimental Psychology: Applied*, 9(4), pp. 219-227.

Favro, D. 2006, "The iconiCITY of ancient Rome", *Urban History*, 33(1), pp. 20-38.

Félonneau, M.L. 2004, "Love and loathing of the city: Urbanophilia and urbanophobia, topological identity and perceived incivilities", *Journal of Environmental Psychology*, 24(1), pp. 43-52.

Feudner, E.M., Engel, C., Neuhann, I.M., Petermeier, K., Bartz-Schmidt, K.U. & Szurman, P. 2009, "Virtual reality training improves wet-lab performance of capsulorhexis: results of a randomized, controlled study*", Graefe's Archive for Clinical and Experimental Ophthalmology*, 247(7), pp. 955-963.

Field, K. & O'Brien, J. 2010, "Cartoblography: Experiments in using and organising the spatial context of micro-blogging", *Transactions in GIS*, 14(SUPPL 1), pp. 5-23.

Fogli, D., Pittarello, F., Celentano, A. & Mussio, P. 2003, “Context-aware interaction in a mobile environment”, *Human-Computer Interaction with Mobile Devices and Services - Lecture Notes in Computer Science*, 2795, pp 434-439

Friedman, M. 1937, “The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance”, *Journal of the American Statistical Association,* 32(200), pp 675-701

Giudice, N.A., Bakdash, J.Z., Legge, G.E. & Roy, R. 2010, "Spatial learning and navigation using: A virtual verbal display", *Transactions on Applied Perception*, 7(1)

Grant, S.C. & Magee, L.E. 1998, "Contributions of proprioception to navigation in virtual environments", *Human Factors*, 40(3), pp. 489-497.

Harrower, M. & Sheesley, B. 2007, "Utterly lost: Methods for reducing disorientation in 3-D fly-over maps", *Cartography and Geographic Information Science*, 34(1), pp. 17-27.

Hasenfratz, J.M., Lapierre, M., Holzschuch, N. & Sillion, F. 2003, "A Survey of Real-time Soft Shadows Algorithms", *Computer Graphics Forum*, 22(4), pp. 753-774.

Hawk, T.F. & Shah, A.J. 2007, "Using Learning Style Instruments to Enhance Student Learning", *Decision Sciences Journal of Innovative Education*, 5(1), pp. 1-19

Hawking, S.W. 1988, *A Brief History of Time: From the Big Bang to Black Holes*, Bantam Press, New York.

Head, D. & Isom, M. 2010, "Age effects on wayfinding and route learning skills", *Behavioural brain research*, 209(1), pp. 49-58.

Hendrix, C. & Barfield, W. 1995, "Relationship between monocular and binocular depth cues for judgements of spatial information and spatial instrument design", *Displays*, 16(3), pp. 103-113.

Hoobler, M., Humphreys, G. & Agrawala, M. 2004, "Visualizing competitive behaviors in multi-user virtual environments", *Proceedings of IEEE Visualization*, VIS 2004, pp. 163-170.

Hubona, G.S., Wheeler, P.N., Shirah, G.W. & Brandt, M. 1999, "The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization", *ACM Transactions on Computer-Human Interaction*, 6(3), pp. 214-242.

Hurlebaus, R., Basten, K., Mallot, H.A. & Wiener, J.M. 2008, “Route learning strategies in a virtual cluttered environment”, *Proceedings of the international conference on Spatial Cognition VI: Learning, Reasoning, and Talking about Space*, pp. 104-120

Huynh, N.T., Hall, G.B., Doherty, S. & Smith, W.W. 2008, "Interpreting urban space through cognitive map sketching and sequence analysis", *Canadian Geographer*, 52(2), pp. 222-240.

Iaria, G., Palermo, L., Committeri, G. & Barton, J.J.S. 2009, "Age differences in the formation and use of cognitive maps*", Behavioural brain research*, 196(2), pp. 187-191.

Ieronutti, L., Ranon, R. & Chittaro, L. 2005, “High-level visualization of users' navigation in virtual environments”, *Proceedings of the 2005 IFIP TC13 international conference on Human-Computer Interaction*, INTERACT'05, pp. 873-885

Imbe, T., Ozaki, F., Kiyasu, S., Mizukami, Y., Ishibashi, S., Inakage, M., Okude, N., Cheok, A.D., Inami, M. & Sugimoto, M. 2010, "Myglobe: A navigation service based on cognitive maps", *Proceedings of the 4th International Conference on Tangible, Embedded, and Embodied Interaction*, TEI'10, pp. 189-192.

Interrante, V., Anderson, L. & Ries, B. 2006, "Distance perception in immersive virtual environments, revisited", *Proceedings of the IEEE conference on Virtual Reality*, VR '06, pp. 3-10.

Interrante, V., Ries, B., Lindquist, J., Kaeding, M. & Anderson, L. 2008, "Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments", *Presence: Teleoperators and Virtual Environments*, 17(2), pp. 176-198.

Jansen-Osmann, P. & Berendt, B. 2002, "Investigating distance knowledge using virtual environments", *Environment and Behavior*, 34(2), pp. 178-193.

Janzen, G., Schade, M., Katz, S. & Herrmann, T. 2001, "Strategies for detour finding in a virtual maze: The role of the visual perspective", *Journal of Environmental Psychology*, 21(2), pp. 149-163.

Jarvis, C.H. 2011, "Spatial Literacy and the Postgraduate GIS Curriculum", *Procedia - Social and Behavioral Sciences*, 21, pp. 294-299.

Jeong, C.-., Chung, Y.-., Joo, S.-. & Lee, J.-. 2006, “Tourism guided information system for location-based services”, *Proceedings of the 2006 international conference on Advanced Web and Network Technologies, and Applications*, APWeb'06, pp. 749-755

Jones, L.A. & Bertamini, M. 2007, "Through the looking glass: How the relationship between an object and its reflection affects the perception of distance and size", *Perception*, 36(11), pp. 1572-1594.

Jurgens, V., Cockburn, A. & Billinghurst, M. 2006, "Depth cues for augmented reality stakeout", *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*, CHINZ ‘06, pp. 117-124.

Kaufman, L. & Kaufman, J.H. 2000, "Explaining the moon illusion", *Proceedings of the National Academy of Sciences of the United States of America,* 97(1), pp. 500-505.

Kellogg, L.H., Bawden, G.W., Bernardin, T., Billen, M., Cowgill, E., Hamann, B., Jadamec, M., Kreylos, O., Staadt, O. & Sumner, D. 2008, "Interactive visualization to advance earthquake simulation", *Pure and Applied Geophysics*, 165(3-4), pp. 621-633.

Kelly, D.M. & Bischof, W.F. 2008, "Orienting in virtual environments: How are surface features and environmental geometry weighted in an orientation task?", *Cognition*, 109(1), pp. 89-104.

Kersten, D., Mamassian, P. & Knill, D.C. 1997, "Moving cast shadows induce apparent motion in depth", *Perception*, 26(2), pp. 171-192.

Kobayashi, T., Miller, H.J. & Othman, W. 2010, "Analytical methods for error propagation in planar space-time prisms", *Journal of Geographical Systems*, 13(4), pp. 327-354.

Koene, R.A., Cannon, R.C. & Hasselmo, M.E. 2003, "Goal-directed spatial navigation of the rat depends on phases of theta oscillation in hippocampal circuitry", *Proceedings of the International Joint Conference on Neural Networks, vol. 2*, pp. 1328-1333.

Koh, R.K.C., Tan, H.X.L. & Duh, H.B.-. 2009, "Information empowerment through mobile learning", *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09

Kühn, P.J. 2004, "Location-based services in mobile communication infrastructures", *AEU - International Journal of Electronics and Communications*, 58(3), pp. 159-164.

Kuijpers, B., Miller, H.J., Neutens, T. & Othmana, W. 2010, "Anchor uncertainty and space-time prisms on road networks", *International Journal of Geographical Information Science*, 24(8), pp. 1223-1248.

Lam, Y.S., Man, D.W.K., Tam, S.F. & Weiss, P.L. 2006, "Virtual reality training for stroke rehabilitation", *NeuroRehabilitation*, 21(3), pp. 245-253.

Lappin, J.S., Shelton, A.L. & Rieser, J.J. 2006, "Environmental context influences visually perceived distance", *Perception and Psychophysics*, 68(4), pp. 571-581.

Lathrop, W.B. & Kaiser, M.K. 2005, "Acquiring spatial knowledge while traveling simple and complex paths with immersive and nonimmersive interfaces", *Presence: Teleoperators and Virtual Environments*,14(3), pp. 249-263.

Lazem, S.Y. & Sheta, W.M. 2005, "Automatic landmark identification in large virtual environment: A spatial data mining approach", *Proceedings of the Ninth International Conference on Information Visualisation*, IV '05, pp. 752-757.

Liu, S., Cheng, D. & Hua, H. 2008, "An optical see-through head mounted display with addressable focal planes", *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, pp. 33-42.

Lucas, J., Thabet, W. & Worlikar, P. 2008, "A VR-based training program for conveyor belt safety", *Electronic Journal of Information Technology in Construction*, 13(13), pp. 381-417.

Lynch, K. 1960, *The Image of the City*, M.I.T., USA.

Meilinger, T. & Knauff, M. 2008, "Ask for directions or use a map: A field experiment on spatial orientation and wayfinding in an urban environment", *Journal of Spatial Science*, 53(2), pp. 13-23.

Micarelli, R. & Pizzioli, G. 2008, "Metropolitan and rural areas: Interscapes as interfaces?", *International Journal of Environmental Research*, 2(1), pp. 1-12.

Milios, E., Kapralos, B., Kopinska, A. & Stergiopoulos, S. 2003, "Sonification of Range Information for 3-D Space Perception", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*,11(4), pp. 416-421.

Miller, J. & Carlson, L. 2011, "Selecting landmarks in novel environments", *Psychonomic Bulletin and Review*, 18(1), pp. 184-191.

Modjeska, D. & Chignell, M. 2003, "Individual differences in exploration using desktop VR", *Journal of the American Society for Information Science and Technology*, 54(3), pp. 216-228.

Morganti, F., Carassa, A. & Geminiani, G. 2007, "Planning optimal paths: A simple assessment of survey spatial knowledge in virtual environments", *Computers in Human Behavior*, 23(4), pp. 1982-1996.

Mueller, S.C., Jackson, C.P.T. & Skelton, R.W. 2008, "Sex differences in a virtual water maze: An eye tracking and pupillometry study", *Behavioural Brain Research*, 193(2), pp. 209-215.

Nakamizo, S. & Imamura, M. 2004, "Verification of Emmert's law in actual and virtual environments", *Journal of Physiological Anthropology and Applied Human Science*, 23(6), pp. 325-329.

Nurminen, A. 2008, "Mobile 3D city maps", *IEEE Computer Graphics and Applications*, 28(4), pp. 20-31.

Omer, I., Goldblatt, R. & Or, U. 2005, "Virtual city design based on urban image theory", *Cartographic Journal*, 42(1), pp. 15-26.

Palermo, L., Iaria, G. & Guariglia, C. 2008, "Mental imagery skills and topographical orientation in humans: A correlation study", *Behavioural Brain Research*, 192(2), pp. 248-253.

Parush, A. & Berman, D. 2004, "Navigation and orientation in 3D user interfaces: The impact of navigation aids and landmarks", *International Journal of Human Computer Studies*, 61(3), pp. 375-395.

Peters, D., Wu, Y. & Winter, S. 2010, “Testing landmark identification theories in virtual environments”, *Proceedings of the 7th international conference on Spatial cognition*, SC'10, pp. 54-69

Philbeck, J.W. & Loomis, J.M. 1997, "Comparison of Two Indicators of Perceived Egocentric Distance under Full-Cue and Reduced-Cue Conditions*", Journal of Experimental Psychology: Human Perception and Performance*, 23(1), pp. 72-85.

Phillips, L., Ries, B., Kaeding, M. & Interrante, V. 2010, "Avatar self-embodiment enhances distance perception accuracy in non-photorealistic immersive virtual environments", *Proceedings of the 2010 IEEE Virtual Reality Conference*, VR '10, pp. 115-118.

Pierce, J.S. & Pausch, R. 2004, "Navigation with place representations and visible landmarks", *Proceedings of the IEEE Virtual Reality 2004*, VR '04, pp. 173-180.

Pires Jr., W.R., Loureiro, A.A.F. & Oliveira, R.A.R. 2010, "Using Web technologies in assessment of context-aware pervasive/ubiquitous systems: A tourist guide service", *2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications*, WiMob'2010, pp. 691-698.

Predebon, J. & Woolley, J.S. 1994, "The familiar-size cue to depth under reduced-cue viewing conditions", *Perception*, 23(11), pp. 1301-1312.

Priestnall, G. & Polmear, G. 2006, "Landscape visualisation: From lab to field", *Proceedings of the First International Workshop on Mobile Geospatial Augmented Reality*

Proffitt, D.R. 2006, "Distance perception", *Current Directions in Psychological Science*, 15(3), pp. 131-135.

Raubal, M. & Winter, S. 2002, "Enriching wayfinding instructions with local landmarks", *Geographic Information Science-Second International Conference,* GIScience 2002, vol. 2478, pp. 243-259.

Rauschecker, A.M., Solomon, S.G. & Glennerster, A. 2006, "Stereo and motion parallax cues in human 3D vision: Can they vanish without a trace?", *Journal of Vision*, 6(12), pp. 1471-1485.

Richardson, T. & Jensen, O.B. 2008, "How mobility systems produce inequality: Making mobile subject types on the Bangkok sky train", *Built Environment*, 34(2), pp. 218-231.

Rousell, A.J., Jarvis, C.H. & Li, J. 2008, "The Effectiveness of Artificial Landmarks for Orientation Tasks Within a Large-Scale Virtual Environment", *Proceedings of the GIS Research UK 16th Annual Conference*, ed. D. Lumbrick, , 2-4 April 2008, pp. 17.

Rowntree, D. 2003, *Statistics without tears: a primer for non-mathematicians (Allyn & Bacon Classics Edition)*, Pearson.

Ruddle, R.A. & Péruch, P. 2004, "Effects of proprioceptive feedback and environmental characteristics on spatial learning in virtual environments", *International Journal of Human Computer Studies*, 60(3), pp. 299-326.

Sadeghian, P., Kantardzic, M., Lozitskiy, O. & Sheta, W. 2006, "The frequent wayfinding-sequence (FWS) methodology: Finding preferred routes in complex virtual environments", *International Journal of Human Computer Studies*, 64(4), pp. 356-374.

Savitch, H.V. 2005, "An anatomy of urban terror: Lessons from Jerusalem and elsewhere", *Urban Studies*, 42(3), pp. 361-395.

Schoor, W., Mecke, R. & Rehfeld, M. 2006, “VR based knowledge transfer in medical technology and techniques”, *Proceedings of the 6th international conference on Computational Science and Its Applications - Volume Part I*, ICCSA'06, pp. 268-277

Shamsuddin, S. & Ujang, N. 2008, "Making places: The role of attachment in creating the sense of place for traditional streets in Malaysia", *Habitat International*, 32(3), pp. 399-409.

Siegel, A.W. & White, S.H. 1975, “The Development of Spatial Representations of Large-Scale Environments”. *Advances in Child Development and Behaviour*, pp. 9-55

Sjölinder, M., Höök, K., Nilsson, L.-. & Andersson, G. 2005, "Age differences and the acquisition of spatial knowledge in a three-dimensional environment: Evaluating the use of an overview map as a navigation aid", *International Journal of Human Computer Studies*, 63(6), pp. 537-564.

Smith, L., Philippides, A. & Husbands, P. 2006, “Navigation in large-scale environments using an augmented model of visual homing”. *Proceedings Animals to Animats 9 – Lecture Notes in Computer Science, vol.4095*, pp. 251-262

Song, Y. & Miller, H.J. 2011, "Exploring traffic flow databases using space-time plots and data cubes", *Transportation*, 39(2), pp. 215-234.

Sorrows, M.E. & Hirtle, S.C. 1999, "The nature of landmarks for real and electronic spaces", *Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*, COSIT '99, pp. 37-50.

Steiner, K.E. & Voruganti, L. 2004, "A comparison of guidance cues in desktop virtual environments", *Virtual Reality*, 7(3-4), pp. 140-147.

Stenton, S.P., Hull, R., Goddi, P.M., Reid, J.E., Clayton, B.J., Melamed, T.J. & Wee, S. 2007, "Mediascapes: Context-aware multimedia experiences", *IEEE Multimedia*, 14(3), pp. 98-105.

Stüdeli, T. 2009, "Surgical wayfinding and navigation processes in the human body", *Cognitive Processing*, 10(suppl 2), pp. 316-318.

Sturz, B.R., Bodily, K.D., Katz, J.S. & Kelly, D.M. 2009, "Evidence against integration of spatial maps in humans: Generality across real and virtual environments", *Animal Cognition*, 12(2), pp. 237-247.

Suma, E.A., Babu, S. & Hodges, L.F. 2007, "Comparison of travel techniques in a complex, multi-level 3D environment", *Proceedings of the IEEE Symposium on 3D User Interfaces 2007,* 3DUI 2007, pp. 147.

Taylor, N. 2003, "The aesthetic experience of traffic in the modern city", *Urban Studies*, 40(8), pp. 1609-1625.

Thompson, W.B., Willemsen, P., Gooch, A.A., Creem-Regehr, S.H., Loomis, J.M. & Beall, A.C. 2004, "Does the quality of the computer graphics matter when judging distances in visually immersive environments?", *Presence: Teleoperators and Virtual Environments*, 13(5), pp. 560-571.

Uehira, K., Suzuki, M. & Abekawa, T. 2007, "3-D display using motion parallax for extended-depth perception", *Proceedings of the 2007 IEEE International Conference on Multimedia and Expo*, ICME 2007, pp. 1742.

Vinson, N.G. 1999, "Design guidelines for landmarks to support navigation in virtual environments", *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '99, pp. 278-285.

Voyer, D., Voyer, S., & Bryden, M.P. (1995) "Magnitude of Sex Differences in Spatial Abilities: A Meta-Analysis and Consideration of Critical Variables", *Psychological Bulletin*, 117(2), pp. 250-270

Wan, G., Jia, F., Liu, Y., Song, G. & Xia, Q. 2008, "Assessing the Availability of Virtual Geographic Environment", *Virtual Geographic Environments*, 7-8 January, 2008.

Watanuki, K. & Kojima, K. 2006, "Virtual reality based knowledge acquisition and job training for advanced casting skills", *Proceedings of the 16th International Conference on Artificial Reality and Telexistence - Workshops*, ICAT 2006, pp. 666.

Willemsen, P. & Gooch, A.A. 2002, "Perceived egocentric distances in real, image-based, and traditional virtual environments", *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, APGV '09, pp. 275.

Willemsen, P., Gooch, A.A., Thompson, W.B. & Creem-Regehr, S.H. 2008, "Effects of stereo viewing conditions on distance perception in virtual environments", *Presence: Teleoperators and Virtual Environments*, 17(1), pp. 91-101.

Witmer, B.G., Bailey, J.H., Knerr, B.W. & Parsons, K.C. 1996, "Virtual spaces and real world places: Transfer of route knowledge", *International Journal of Human Computer Studies*, 45(4), pp. 413-428.

Witmer, B.G. & Kline, P.B. 1998, "Judging perceived and traversed distance in virtual environments", *Presence: Teleoperators and Virtual Environments*, 7(2), pp. 144-167.

Witmer, B.G., Sadowski, W.J. & Finkelstein, N.M. 2002, "VE-based training strategies for acquiring survey knowledge", *Presence: Teleoperators and Virtual Environments*, 11(1), pp. 1-18.

Woollett, K. & Maguire, E.A. 2010, "The effect of navigational expertise on wayfinding in new environments", *Journal of Environmental Psychology*, 30(4), pp. 565-573.

Wu, A., Zhang, W. & Zhang, X. 2009, "Evaluation of wayfinding aids in virtual environment", *International Journal of Human-Computer Interaction*, 25(1), pp. 1-21.

Wu, B., He, Z.J. & Ooi, T.L. 2007a, "Inaccurate representation of the ground surface beyond a texture boundary", *Perception*, 36(5), pp. 703-721.

Wu, B., He, Z.J. & Ooi, T.L. 2007b, "The linear perspective information in ground surface representation and distance judgment", *Perception and Psychophysics*, 69(5), pp. 654-672.

Yaegashi, K. & Yanai, K. 2009, Can geotags help image recognition?, *Proceedings of the 3rd Pacific Rim Symposium on Advances in Image and Video Technology*, PSIVT '09, pp. 361-373

Yang, H., An, X., Wu, J. & Tang, G. 2010, "Autonomous navigation based on stellar refraction for elliptic orbit spacecraft", 3rd International Symposium on Systems and Control in Aeronautics and Astronautics (ISSCAA), pp. 1019-1022.

Zanbaka, C.A., Lok, B.C., Babu, S.V., Ulinski, A.C. & Hodges, L.F. 2005, "Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment", *IEEE Transactions on Visualization and Computer Graphics*, 11(6), pp. 694-705.

Zhang, X. 2008, "A multiscale progressive model on virtual navigation", *International Journal of Human Computer Studies,* 66(4), pp. 243-256.

Map data (roads, rivers, elevation etc.) in figures 5, 6, 7, 8, 9, 15, 16, 21, 31, 41, 46, 47, 48, 54, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69 from the Spatial Data Infrastructure of Andalusia (Junta de Andalucía).