

The effect of 2D vs. 3D visualisation on lidar point cloud analysis tasks

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

by

**Claire Leonora Burwell
Department of Geography
University of Leicester**

2016

The effect of 2D vs. 3D visualisation on lidar point cloud analysis tasks

Claire Leonora Burwell

Abstract

The exploitation of human depth perception is not uncommon in visual analysis of data; medical imagery and geological analysis already rely on stereoscopic 3D visualisation. In contrast, 3D scans of the environment are usually represented on a flat, 2D computer screen, although there is potential to take advantage of both (a) the spatial depth that is offered by the point cloud data, and (b) our ability to see stereoscopically. This study explores whether a stereo 3D analysis environment would add value to visual lidar tasks, compared to the standard 2D display.

Forty-six volunteers, all with good stereovision and varying lidar knowledge, viewed lidar data in either 2D or in 3D, on a 4m x 2.4m screen. The first task required 2D and 3D measurement of linear lengths of a planar and a volumetric feature, using an interaction device for point selection. Overall, there was no significant difference in the spread of 2D and 3D measurement distributions for both of the measured features.

The second task required interpretation of ten features from individual points. These were highlighted across two areas of interest - a flat, suburban area and a valley slope with a mixture of features. No classification categories were offered to the participant and answers were expressed verbally. Two of the ten features (chimney and cliff-face) were interpreted with a better degree of accuracy using the 3D method and the remaining features had no difference in 2D and 3D accuracy.

Using the experiment's data processing and visualisation approaches, results suggest that stereo 3D perception of lidar data does not add value to manual linear measurement. The interpretation results indicate that immersive stereo 3D visualisation does improve the accuracy of manual point cloud classification for certain features. The findings contribute to wider discussions in lidar processing, geovisualisation, and applied psychology.

Acknowledgements

Firstly, thank you to all of the volunteers who gave up their time to take part in the pilot study and the experiment.

I thank my supervisors, Claire Jarvis and Kevin Tansey, for the opportunity to develop my MSc research and delve deeper into the world of lidar, stereovision, and participant studies. I appreciate their support and guidance throughout my PhD research, which was partially funded by the University of Leicester's (UoL) College of Science & Engineering's PhD scholarship scheme and the University's Research Training Support Grant.

Special thanks go to Ned Chisholm, of Airbus Defence and Space Ltd., for sharing his time and lidar expertise. I am grateful too to Phil Duke, from the UoL Department of Psychology, for a key discussion about human vision. Within the UoL Geography Department, thanks to Computer Officer, Adam Cox, for help during my experiment development when I came across various IT challenges. I am appreciative of the everyday consideration of colleagues during the trials, between Oct 2013 – Feb 2014, especially from Jen Dickie, Laura Wilson, and Gary Hancox, which helped the smooth-running of the experiments. Thanks too to Cartographer Kerry Allen for the Illustrator training, which I put to good use.

I am thankful for the camaraderie of my fellow PhD colleagues over the years, thanks in particular to Paul Arellano, Firdos Almadani, Mustafa Kose, David Ackerley, and Hannah Brooking. Thanks to former PhD students from the department who passed on invaluable tidbits of advice, especially fellow VR-student Adam Rousell. Big thanks to former UoL Research Assistants Sarah Mills, Alberto Ramirez, Joe Dutton, and Beth Cole for their friendship and motivation. I am also grateful to Ed Manley and Faye Outram for sharing their PhD pros and cons prior to starting my doctorate.

Thanks to those away from uni who have been encouraging me throughout the doctoral journey - Chloë Brown, Emily Kenrick, and Emma Casson. I am grateful for the backing and help from my parents, Pete and Norm. Thank you especially to Matt Driver – *merci Mathieu!*

List of Contents

| | |
|---|------------|
| Abstract | i |
| Acknowledgements | ii |
| List of Contents | iii |
| List of Figures | vi |
| List of Tables | xi |
| List of Equations | xiv |
| Glossary | xv |
| 1. Introduction | 2 |
| 1.1 RESEARCH AIM | 7 |
| 1.2 THESIS STRUCTURE | 8 |
| 2. Literature review | 11 |
| 2.1 INTRODUCTION | 11 |
| 2.2 LIDAR POINT CLOUD VISUALISATION | 11 |
| 2.2.1 BACKGROUND TO LIDAR | 11 |
| 2.2.2 IMPORTANCE OF POINT CLOUD VISUALISATION | 15 |
| 2.2.3 STANDARD 2D VISUALISATION | 20 |
| 2.2.4 PROPOSED 3D VISUALISATION | 23 |
| 2.3 EVALUATING GEOVISUALISATION METHODS | 29 |
| 2.3.1 PARTICIPANT EVALUATION | 30 |
| 2.3.2 EXISTING VS. PROTOTYPE | 31 |
| 2.4 REVIEW OUTCOME | 32 |
| 2.4.1 RESEARCH QUESTIONS & HYPOTHESES | 33 |
| 3. Method | 36 |
| 3.1 INTRODUCTION | 36 |
| 3.2 DEVELOPMENT | 36 |
| 3.2.1 LIDAR DATA PREPARATION | 36 |
| 3.2.2 VISUALISATION SYSTEM DEVELOPMENT | 48 |
| 3.2.3 EXPERIMENT DESIGN | 54 |
| 3.3 PILOT STUDY | 57 |
| 3.3.1 SET-UP OF PILOT | 58 |
| 3.3.2 PILOT PARTICIPANTS | 59 |

| | | |
|------------|---|------------|
| 3.3.3 | EVALUATING SUITABILITY..... | 60 |
| 3.3.4 | AMENDMENTS TO THE EXPERIMENT..... | 64 |
| 3.4 | MAIN EXPERIMENT _____ | 65 |
| 3.4.1 | RECRUITMENT CAMPAIGN..... | 66 |
| 3.4.2 | EXPERIMENT SET-UP..... | 67 |
| 3.4.3 | EXPERIMENT STAGES..... | 69 |
| 3.4.4 | DATA OUTPUTS..... | 73 |
| 3.5 | DATA PROCESSING & ANALYSIS _____ | 74 |
| 3.5.1 | EXPERIMENT DATA CONVERSION AND COLLATION..... | 74 |
| 3.5.2 | DATA ANALYSIS..... | 76 |
| 3.6 | SUMMARY _____ | 77 |
| 4. | Participant background _____ | 80 |
| 4.1 | INTRODUCTION _____ | 80 |
| 4.2 | METHOD _____ | 80 |
| 4.2.1 | DEMOGRAPHIC DATA..... | 80 |
| 4.2.2 | RANKING PARTICIPANT LIDAR EXPERTISE..... | 80 |
| 4.3 | CHARACTERISTICS OF OVERALL SAMPLE POPULATION _____ | 82 |
| 4.3.1 | DEMOGRAPHICS..... | 82 |
| 4.4 | SCREENING OF PARTICIPANTS _____ | 86 |
| 4.4.1 | STEREOACUITY RESULTS..... | 86 |
| 4.5 | CHARACTERISTICS PER TRIAL GROUP _____ | 89 |
| 4.5.1 | DEMOGRAPHICS PER 2D/3D GROUP..... | 90 |
| 4.5.2 | A PRIORI LIDAR KNOWLEDGE AND EXPERIENCE PER 2D/3D GROUP..... | 93 |
| 4.5.3 | TECHNOLOGY HABITS PER 2D/3D GROUP..... | 94 |
| 4.6 | SUMMARY _____ | 96 |
| 5. | Measurement _____ | 102 |
| 5.1 | INTRODUCTION _____ | 102 |
| 5.2 | METHOD _____ | 103 |
| 5.2.1 | TASK DEVELOPMENT..... | 103 |
| 5.2.2 | EXPERIMENT INSTRUCTIONS..... | 108 |
| 5.2.3 | DATA ANALYSIS..... | 112 |
| 5.3 | RESULTS _____ | 116 |
| 5.3.1 | PLANAR FEATURE (SCENE A) – ROOF EDGE LENGTH..... | 116 |
| 5.3.2 | VOLUMETRIC FEATURE (SCENE B) - CANOPY DIAMETER ESTIMATES..... | 124 |
| 5.3.3 | RESULTS SUMMARY..... | 132 |
| 5.4 | DISCUSSION _____ | 133 |
| 5.4.1 | 2D VS. 3D PLANAR MEASUREMENT PRECISION (RQ1.2A)..... | 133 |
| 5.4.2 | 2D VS. 3D VOLUMETRIC MEASUREMENT PRECISION (RQ1.2B)..... | 134 |
| 5.5 | REFLECTION ON METHODOLOGICAL DEVELOPMENT (RQ3) _____ | 136 |
| 5.5.1 | INTERACTION TECHNIQUE..... | 137 |
| 5.5.2 | VEGETATION ANALYSIS APPROACH..... | 140 |
| 5.5.3 | 2D VS. 3D ASSESSMENT METHODS..... | 141 |
| 5.6 | SUMMARY _____ | 141 |

6. Interpretation 143

| | |
|---|------------|
| 6.1 INTRODUCTION | 143 |
| 6.2 METHOD | 144 |
| 6.2.1 TASK DEVELOPMENT | 144 |
| 6.2.2 EXPERIMENT INSTRUCTIONS | 154 |
| 6.2.3 DATA ANALYSIS | 158 |
| 6.3 RESULTS | 164 |
| 6.3.1 2D VS. 3D INTERPRETATION OF FEATURES (POIS A – J) | 165 |
| 6.3.2 EFFECT OF PHYSICAL AOI (SCENES C & D)..... | 172 |
| 6.3.3 RESULTS SUMMARY | 175 |
| 6.4 DISCUSSION | 177 |
| 6.4.1 2D VS. 3D INTERPRETATION OF DIFFERENT FEATURES (RQ2.1)..... | 177 |
| 6.4.2 2D VS. 3D INTERPRETATION OF DIFFERENT ENVIRONMENTS (RQ 2.2) | 181 |
| 6.5 REFLECTION ON METHODOLOGICAL DEVELOPMENT (RQ3) | 183 |
| 6.5.1 SELECTION OF POIS | 183 |
| 6.5.2 RESTRICTION OF LEARNING | 184 |
| 6.5.3 CLASSIFICATION TECHNIQUE | 185 |
| 6.6 SUMMARY | 187 |

7. Method Review 190

| | |
|--|------------|
| 7.1 INTRODUCTION | 190 |
| 7.2 REFLECTION ON GENERAL METHOD DEVELOPMENT (RQ3) | 190 |
| 7.2.1 DEPTH PERCEPTION..... | 190 |
| 7.2.2 HUMAN FACTORS | 191 |
| 7.2.3 EFFECT OF PILOT STUDY | 191 |
| 7.2.4 QUALITATIVE DATA..... | 192 |
| 7.3 METHOD RECOMMENDATIONS | 193 |
| 7.4 FUTURE WORK | 194 |
| 7.4.1 2D VS. 3D EVALUATION | 194 |
| 7.4.2 2D VS. 3D HUMAN COMPUTER INTERACTION METHODS | 195 |
| 7.4.3 EFFECT OF LIDAR DATA REPRESENTATION ON 2D VS. 3D RESULTS | 196 |
| 7.5 SUMMARY | 198 |

8. Conclusion 200

| | |
|---|------------|
| 8.1 INTRODUCTION | 200 |
| 8.2 FINDINGS | 200 |
| 8.2.1 MEASUREMENT TASK (RQ1)..... | 200 |
| 8.2.2 INTERPRETATION TASK (RQ2) | 201 |
| 8.2.3 REFLECTION ON METHODOLOGICAL DEVELOPMENT (RQ3)..... | 202 |
| 8.3 IMPACT | 203 |
| 8.4 CONCLUSION | 204 |

Bibliography 207

Appendices 221

List of Figures

| | |
|---|----|
| FIGURE 1-1. ILLUSTRATION OF THE HUMAN FIELD OF VISION, SHOWN FROM PLAN VIEW..... | 2 |
| FIGURE 1-2. THE VISUAL PERCEPTION OF OBJECTS IN 3D SPACE, COMPARED TO THEIR PICTORIAL REPRESENTATION ON A FLAT SCREEN. OVAL REPRESENTS HUMAN HEAD. ADAPTED FROM HODGES (1992, CITED IN HOWARD AND ROGERS, 2012, P.539)..... | 3 |
| FIGURE 1-3. SUMMARY OF VISUAL DEPTH CUES. ADDITIONAL PICTORIAL CUES ARE DESCRIBED IN TOVÉE (1996)..... | 4 |
| FIGURE 1-4. SCHEMA SHOWING THE PROCESS OF VISUALISATION (ADAPTED FROM WARE, 2012)..... | 5 |
| FIGURE 1-5. GROUND TRUTH IMAGE (GOOGLE, 2014), LEFT, DEPICTING SUBURBAN FEATURE AND, RIGHT, ITS EQUIVALENT LASER-SCANNED POINT CLOUD. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A). DATA WERE CAPTURED USING AN OPTTECH GEMINI AIRBORNE SENSOR AND INCLUDED OVERLAPPING FLIGHTLINES (SEE APPENDIX A) WITH ~8 POINTS PER METRE SQUARED (PPM ²). RIGHT-HAND IMAGE SHOWS PROCESSED DATA, WITH BUILDING POINTS AT A DENSITY OF ~7PPM ² AND GROUND POINTS AT ~2PPM ² . THE FULL LIDAR PROCESSING METHOD IS DETAILED IN CHAPTER 3. | 6 |
| FIGURE 1-6. THESIS STRUCTURE. BOLD ELEMENTS DENOTE RESEARCH QUESTIONS (RQ)..... | 8 |
| FIGURE 2-1. IMAGE TO SHOW DATA ACQUISITION EXAMPLE USING LASER-SCANNER MOUNTED IN A PLANE, SCANNING AN URBAN GEOGRAPHICAL FEATURE WITHIN THE 3D PHYSICAL ENVIRONMENT. | 12 |
| FIGURE 2-2. ILLUSTRATION OF LASER-SCANNING DATA ACQUISITION TECHNIQUES. IMAGE NOT TO (RELATIVE) SCALE. FROM LEFT TO RIGHT, LASER SENSOR MOUNTED ON SATELLITE, AIRCRAFT/HELICOPTER, DRONE (UNMANNED AERIAL VEHICLE), ROAD VEHICLE, TRIPOD. DATA CAPTURE METHODS VARY BETWEEN INSTRUMENTS. | 13 |
| FIGURE 2-3. ILLUSTRATION DEPICTING A VOLUME OF RAW POINT CLOUD (LEFT), WHICH IS MADE UP OF XYZ POINTS COLLECTED VIA AIRBORNE LASER-SCANNING. THE RAW DATA POINTS CAN BE SEGMENTED INTO DIFFERENT CLASSIFICATIONS (LEFT), E.G. GROUND SURFACE AND VEGETATION. DIAGRAM BASED ON LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A). DISPLAYED DENSITY IS ~2 POINTS PER METRE SQUARED (PPM ²) FOR GROUND POINTS AND ~4PPM ² FOR VEGETATION. | 14 |
| FIGURE 2-4. FLOW DIAGRAM ADAPTED FROM FLOOD (2001) SHOWING THE ACTIONS REQUIRED DURING A COMMERCIAL LIDAR DATA WORKFLOW, BEFORE DELIVERY OF THE DATASET TO THE CLIENT. ASTERISKS HIGHLIGHT THE HUMAN ASPECTS OF THE PROCESS, DURING WHICH 60-80% OF PRODUCTION OCCURS. | 16 |
| FIGURE 2-5. FROM LIDAR DATA ACQUISITION OF THE 3D PHYSICAL ENVIRONMENT, TO 2D VISUALISATION OF THE DATA. THE 3D DATA ARE FLATTENED – DOES THIS AFFECT VISUALISATION OUTPUTS? DIAGRAM AUTHOR'S OWN. | 22 |
| FIGURE 2-6. FROM LIDAR DATA ACQUISITION OF THE 3D PHYSICAL ENVIRONMENT, TO STEREO 3D VISUALISATION OF THE DATA. COULD THE DATA OUTPUT BE AN IMPROVEMENT ON 2D DISPLAY-DERIVED DATA? DIAGRAM AUTHOR'S OWN..... | 24 |
| FIGURE 2-7. DIFFERENT STAGES OF IMMERSION, AFTER BRODLIE ET AL. (2002) | 25 |
| FIGURE 2-8. THREE COGNITIVE SPACES, ADAPTED FROM MARK (1992). M REPRESENTS THE 'METAPHORS OR MAPPINGS' BETWEEN THE SPACES. TRANSPERCEPTUAL SPACE IS BUILT FROM THE EXPERIENCE OF OTHER SPACES. | 26 |

| | |
|--|----|
| FIGURE 2-9. VISUAL SUMMARY OF PROBLEM STATEMENT – WHICH IS MORE PRECISE/ACCURATE, WITH RESPECT TO LIDAR POINT CLOUDS, DATA OUTPUT FROM 2D OR STEREOSCOPIC 3D VISUALISATION? HERE, 2D REFERS TO 2.5D REPRESENTATION OF (2D OR 3D) OBJECTS ON A FLAT PROJECTION, WHEREAS OFFSET PROJECTIONS OF THE SAME 2D/3D OBJECT TO EACH EYE RESULT IN VIEWER VISUALISING THE OBJECT IN 3D SPACE. | 33 |
| FIGURE 3-1. TIMELINE OF THE DIFFERENT STAGES OF THE STUDY’S METHODOLOGY, FOLLOWING THE INITIAL YEAR OF LITERATURE REVIEW AND PROJECT DEVELOPMENT. | 36 |
| FIGURE 3-2. NODE TO SHOW THE 3 ASPECTS OF EXPERIMENT DEVELOPMENT, WITH EMPHASIS ON THE FIRST LIDAR DATA PREPARATION STAGE. | 37 |
| FIGURE 3-3. MAP SHOWING THE LOCATIONS OF THE LASER-SCANNING ACQUISITION SITES IN BRISTOL AND LONDON, UK, WITHIN ORDNANCE SURVEY (OS) NATIONAL GRID TILES ST AND TQ. ORDNANCE SURVEY DATA © CROWN COPYRIGHT AND DATABASE RIGHT 2014. | 38 |
| FIGURE 3-4. FLOWLINE OF LIDAR DATA PROCESSING AND THE SOFTWARE (DETAILED IN TABLE 1-2) USED TO CARRY OUT EACH STAGE. ACQUISITION AND PRE-PROCESSING STAGES WERE CARRIED OUT BY AIRBUS DEFENCE AND SPACE LTD., DURING 3D URBAN MODELLING PRODUCT GENERATION FOR UK CITIES, IN 2008. THE REMAINING PROCESSING WAS UNDERTAKEN DURING THIS RESEARCH PROJECT, IN 2012. | 40 |
| FIGURE 3-5. SCHEMA SHOWING THE POSITIONS AT WHICH AIRBORNE LIDAR RETURNS OCCUR. THE SOLD LINE DENOTES AN IN-COMING LASER THAT HAS BEEN FIRED FROM AN AIRBORNE INSTRUMENT. IMAGE AUTHOR’S OWN. | 42 |
| FIGURE 3-6. DIAGRAM HIGHLIGHTING THE CONSIDERATIONS OF VISUALISATION SYSTEM DEVELOPMENT. | 48 |
| FIGURE 3-7. HARDWARE USED BY PARTICIPANTS IN THE EXPERIMENT. SAITEK GAMEPAD (SAITEK 2007), LEFT, AND XPAND X102 ACRIVE SHUTTER 3D GLASSES (WWW.XPAND.ME, ACCESSED 01-04-15). IMAGES AUTHOR’S OWN. | 49 |
| FIGURE 3-8. SCREENSHOT FROM TERRASCAN SOFTWARE (TERRASOLID, 2013), NO SCALE, SHOWING BRISTOL LIDAR DATA POINTS (ACQUIRED 25-11-2008) OVERLYING ORTHOPHOTOGRAPHY (14-04-2007). ARROWS HIGHLIGHT THE IMPACT OF AERIAL IMAGERY LEAN ON TRUE-COLOURING OF LIDAR POINT CLOUD. RED POINTS = BUILDING, GREEN = VEGETATION, ORANGE = GROUND. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTOGRAPH © GEOPERSPECTIVES (2013). | 53 |
| FIGURE 3-9. DIAGRAM HIGHLIGHTING THE EXPERIMENT DESIGN ASPECT OF THE EXPERIMENT DEVELOPMENT. | 55 |
| FIGURE 3-10. RANDOT™ ('RANDOM DOT') STEREO TEST (STEREO OPTICAL Co., 2009). THE 'CIRCLES (WITH RANDOM DOT GROUND)' TEST, WAS CARRIED OUT BY EACH PARTICIPANT TO DETERMINE THEIR STEREOACUITY. 10/10 CORRECT ANSWERS = 400 SECONDS OF ARC. AS SHOWN IN THE IMAGE, THE OTHER TESTS WERE OBSCURED BY CARDBOARD. IMAGE AUTHOR’S OWN. | 56 |
| FIGURE 3-11. ELEMENTS OF THE PILOT STAGE OF THE METHODOLOGY. | 58 |
| FIGURE 3-12. PLOTS SHOWING PILOT PARTICIPANTS’ RANKING OF DIFFERENT COLOUR REPRESENTATIONS OF LIDAR POINT CLOUD DATA – ONE COLOUR (GREEN POINTS WITH BLACK BACKGROUND), GREYSCALE (BY Z VALUE), THEMATIC RGB COLOURING ACCORDING TO CLASSIFICATION. X AXIS RANKS EACH COLOUR REPRESENTATION FROM 1-7, WITH 1 BEING VERY EFFECTIVE AT HELPING THEM UNDERSTAND THE STRUCTURE AND 7 BEING NOT AT ALL EFFECTIVE. TOP ROW AND MIDDLE ROW SHOW FEEDBACK FROM PILOT2, BOTTOM ROW SHOWS RESULTS FROM PILOT3 AND PILOT4 REGARDING GENERAL DISPLAY (I.E. REGARDLESS OF 2D OR 3D METHOD). NOTE DIFFERENT Y AXIS IN ONE COLOUR GENERAL DISPLAY. | 63 |

| | |
|--|----|
| FIGURE 3-13. FLOW DIAGRAM HIGHLIGHTING THE TIMELINE FOR THE MAIN EXPERIMENT STAGE OF THE METHODOLOGY | 66 |
| FIGURE 3-14. DIFFERENT ORDERS, RELATING TO THE COMBINATIONS OF SCENES (A, B, C, AND D) AND VISUALISATION METHODS (2D OR 3D) THAT WERE RANDOMLY ASSIGNED TO PARTICIPANTS. | 67 |
| FIGURE 3-15. SCHEMA (LEFT) SHOWING THE LAYOUT OF THE VIRTUAL REALITY THEATRE WITHIN THE GEOGRAPHY DEPT AT THE UNIVERSITY OF LEICESTER, WHICH WAS USED DURING TRIALS. THE ROOM IS APPROXIMATELY 6M X 7M. THE CHEQUERBOARD ELEMENTS REPRESENT THE 3D STEREO HARDWARE, WHICH IS MOUNTED ON THE CEILING. ALL OTHER GREY ELEMENTS ARE FURNITURE, DESKS OR CHAIRS, THAT WERE PRESENT, BUT NOT REQUIRED. DASHED LINE INDICATES AREA FOR INTRODUCTORY DISCUSSIONS AND FEEDBACK, PRIOR TO THE EXPERIMENT. TOP PHOTO SHOWS THE SET-UP, WHERE R = RESEARCHER'S POSITION AND P = PARTICIPANT. | 68 |
| FIGURE 3-16. EXPERIMENT FLOW - THE PARTICIPANTS WHO PASSED THE STEREO SCREENING TEST AND CONTINUED TO CARRY OUT THE MAIN EXPERIMENT. | 70 |
| FIGURE 3-17. TWO SLIDES PRESENTED TO THE PARTICIPANTS TO EXPLAIN THE CONCEPT OF LIDAR DATA ACQUISITION FOR MAN-MADE FEATURES (TOP) AND NATURAL FEATURES (BOTTOM). IMAGES AUTHOR'S OWN, PRODUCED USING MICROSOFT POWERPOINT (MICROSOFT CORPORATION, 2010). | 71 |
| FIGURE 3-18. SUMMARY OF PARTICIPANT GUIDELINES FOR GAMEPAD TRAINING. LEFT ANALOGUE STICK = MOVEMENT, RIGHT ANALOGUE STICK = HEAD TURN/LOOK..... | 72 |
| FIGURE 3-19. DIAGRAM OF THE DATA PROCESSING STAGE – DATA CONVERSION, COLLATION AND ANALYSIS. THE ANALYSIS METHODS USED FOR EACH TASK ARE DETAILED IN CHAPTERS 5 AND 6. | 74 |
| FIGURE 3-20. SCHEMA SHOWING A NORMAL FREQUENCY DISTRIBUTION (IN PERCENT) OF A SAMPLE AGAINST A VARIABLE. SHADED PORTIONS UNDER THE BELL-SHAPED CURVE SHOW THE 5% CRITICAL REGION AT WHICH THE NULL HYPOTHESIS (H_0) IS REJECTED. IMAGE AUTHOR'S OWN. | 77 |
| FIGURE 3-21. SCHEMA TO SHOW ELEMENTS OF A BOX PLOT, IN RELATION TO THE EQUIVALENT HISTOGRAM. ABSENCE OF THE IQR BOX INDICATES $\geq 50\%$ OF THE SAMPLE POPULATION EXHIBITED THE SAME VALUE. IMAGE AUTHOR'S OWN..... | 77 |
| FIGURE 3-22. A SUMMARY OF THE METHOD - NODES RELATE TO THE DIFFERENT STAGES OF PROJECT DEVELOPMENT. | 78 |
| FIGURE 4-1. GENDER SPLIT OF VOLUNTEERS..... | 83 |
| FIGURE 4-2. DISTRIBUTION OF AGE RANGES OF VOLUNTEERS. | 83 |
| FIGURE 4-3. GRAPHS DESCRIBING THE OCCUPATION OF THE SAMPLE POPULATION. INSET GRAPH SHOWS THE LEVEL OF DEGREES BEING PURSUED BY THE STUDENT CATEGORY..... | 84 |
| FIGURE 4-4. NATIONALITY OF ALL PARTICIPANTS..... | 85 |
| FIGURE 4-5. NATIVE LANGUAGE OF ALL PARTICIPANTS. | 85 |
| FIGURE 4-6. PARTICIPANT MOTIVATION TO TAKE PART IN THE STUDY. | 86 |
| FIGURE 4-7. BAR GRAPH SHOWING THE LEVELS OF STEREOACUITY REACHED BY THE GENERAL SAMPLE OF PARTICIPANTS (N= 51). RESULTS OF 6/10 OR ABOVE MEANT THAT PARTICIPANTS WERE NOT ALLOWED TO TAKE PART IN THE FULL EXPERIMENT. 6/10 OR HIGHER IS THE EQUIVALENT OF 50 SECONDS OF ARC OR LOWER..... | 87 |

| | |
|--|-----|
| FIGURE 4-8. BOXPLOTS SHOWING THE DISTRIBUTION OF STEREO SCORES ACHIEVED BY FEMALES AND MALES WHO (LEFT) VOLUNTEERED TO TAKE PART IN THE EXPERIMENT AND (RIGHT) PASSED THE STEREO TEST AND TOOK PART IN THE MAIN EXPERIMENT. DARK LINE DENOTES MEDIAN. | 88 |
| FIGURE 4-9. AGE-RANGES OF PARTICIPANTS FOR SCENES A TO D, FOR 2D (TOP) AND 3D GROUPS (BOTTOM). | 90 |
| FIGURE 4-10. BARCHARTS SHOWING GENDER SPLIT OF PARTICIPANT GROUPS FOR SCENES A TO D, FOR 2D (TOP) AND 3D GROUPS (BOTTOM). | 91 |
| FIGURE 4-11. BOXPLOTS SHOWING FREQUENCY DISTRIBUTION OF PARTICIPANT STEREOACUITY FOR SCENES A TO D, FOR 2D (TOP) AND 3D GROUPS (BOTTOM). NUMBER OF PARTICIPANTS ARE PLOTTED AGAINST RANDOT TEST SCORES 6/10 TO 10/10 (Y AXIS). | 92 |
| FIGURE 4-12. 2D PARTICIPANTS' LEVELS OF LIDAR/LASER-SCANNING KNOWLEDGE (TOP) AND EXPERIENCE (BOTTOM). FOR KNOWLEDGE, 0 = NOVICE, 1 = INFORMED, 2 = EXPERT. FOR EXPERIENCE, 0 = NOVICE, 1 = INFORMED, 2 = EXPERT. | 93 |
| FIGURE 4-13. 3D PARTICIPANTS' LEVELS OF LIDAR/LASER-SCANNING KNOWLEDGE (TOP) AND EXPERIENCE (BOTTOM). FOR KNOWLEDGE, 0 = NOVICE, 1 = INFORMED, 2 = EXPERT. FOR EXPERIENCE, 0 = NOVICE, 1 = INFORMED, 2 = EXPERT. | 94 |
| FIGURE 4-14. SUMMARY OF FREQUENCY OF NAVIGATION DEVICE USAGE, FOR ALL PARTICIPANTS. | 95 |
| FIGURE 4-15. FREQUENCY THAT 2D (TOP) AND 3D (BOTTOM) GROUPS OF PARTICIPANTS, FOR EACH SCENE, USE A GAMEPAD DEVICE DURING COMPUTER GAMING. | 97 |
| FIGURE 4-16. SUMMARY OF FREQUENCY OF 3D DISPLAY USAGE, FOR ALL PARTICIPANTS. | 98 |
| FIGURE 4-17. FREQUENCY THAT 2D GROUPS OF PARTICIPANTS, FOR EACH SCENE, EXPERIENCE 3DTV OR 3D CINEMA SCREEN (FILM, SPORTING EVENT, ETC.). *'EFFECTIVE', MEANING YOU FELT THAT THE 3D EXPERIENCE GAVE ADDED DEPTH TO THE IMAGES. RESPONSE RANKED FROM 1, NOT AT ALL EFFECTIVE (NOT EXTRA DEPTH IN 3D), TO 5, VERY EFFECTIVE (VERY STRONG 3D DEPTH). | 99 |
| FIGURE 4-18. FREQUENCY THAT 3D GROUPS OF PARTICIPANTS, FOR EACH SCENE, EXPERIENCE 3DTV OR 3D CINEMA SCREEN (FILM, SPORTING EVENT, ETC.). *'EFFECTIVE', MEANING YOU FELT THAT THE 3D EXPERIENCE GAVE ADDED DEPTH TO THE IMAGES. RESPONSE RANKED FROM 1, NOT AT ALL EFFECTIVE (NOT EXTRA DEPTH IN 3D), TO 5, VERY EFFECTIVE (VERY STRONG 3D DEPTH). | 100 |
| FIGURE 5-1. PLAN AND SIDE VIEWS OF SCENE A POINT CLOUD, WHICH MEASURES (WIDTH X LENGTH X HEIGHT) 29.78M X 33.7M X 15.1M AND IS MADE UP OF 824 GROUND POINTS AND 1323 BUILDING POINTS. AIRBORNE LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A) | 104 |
| FIGURE 5-2. PLAN AND SIDE VIEWS OF SCENE B, WHICH MEASURES (WIDTH X LENGTH X HEIGHT) APPROX. 44M X 38M X 25M. THE POINTS CLOUD IS MADE UP OF 1699 GROUND POINTS AND 2666 HIGH VEGETATION POINTS. AIRBORNE LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A). | 105 |
| FIGURE 5-3. ISSUES WITH ALS HEIGHT MEASUREMENT, DIRECTLY FROM POINT CLOUD. LEFT - LIDAR POINTS AVAILABLE FOR MEASURING HEIGHT OF BUILDING (SOLID VERTICAL ARROW), MIDDLE - MEASURING FROM POINT ABOVE GROUND POINT GIVES AN UNDERESTIMATED HEIGHT, RIGHT - MEASURING FROM THE CORRECT HEIGHT, AT A DIFFERENT XY POSITION LEADS TO OVERESTIMATION OF THE FEATURE HEIGHT. AUTHOR'S OWN IMAGE. | 107 |
| FIGURE 5-4. ELEMENTS OF THE MEASUREMENT TASK IN RELATION TO FIRST AND SECOND SCENES. VERTICAL ARROW INDICATES PASSAGE OF TIME. BLACK TRIANGLES DENOTE FILE OUTPUTS FROM THE TASK. | 108 |

| | |
|--|-----|
| FIGURE 5-5. POINT SELECTION TRAINING – GUIDELINES SHOWN TO PARTICIPANTS, INDICATING GAMEPAD BUTTONS USED TO TOGGLE ON/OFF CROSSHAIR TARGET ICON AND SELECT 1ST AND 2ND POINTS..... | 109 |
| FIGURE 5-6. PHOTOGRAPHS OF EXPERIMENT SET-UP DURING MEASUREMENT TASK. BOXED IMAGES REPRESENT ACTION CARRIED OUT BY PARTICIPANT. OTHER IMAGES SHOW THE RESULTING VECTOR THAT IS GENERATED BETWEEN THE PARTICIPANTS' TWO POINTS – IN SITU AND SCREENSHOT EXAMPLES. PARTICIPANT POSED BY MODEL. AIRBORNE LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A)..... | 111 |
| FIGURE 5-7. SCHEMAS COMPARING PRECISION AND ACCURACY OF OBSERVED VALUES USING A DART BOARD AND THE APPROXIMATE EQUIVALENT HISTOGRAM SHAPES (WHERE X AXIS IS THE OBSERVED MEASUREMENT VALUES AND Y AXIS IS THE FREQUENCY DISTRIBUTION OF THE SAMPLE GROUP). ALS DATA WERE ACQUIRED IN 2008, SO TIME OF FLIGHT GROUND SURVEY WAS NOT AVAILABLE FOR AN APPROPRIATE REFERENCE VALUE. AUTHOR'S OWN DIAGRAM..... | 113 |
| FIGURE 5-8. EXAMPLE OF A Q-Q PLOT SHOWING PARTICIPANT MEASUREMENTS (DOTS) RELATION TO NORMAL DISTRIBUTION. Y AXIS = MEASUREMENT VALUES (M); IF ALL PARTICIPANTS MEASURE THE SAME DISTANCE, THE DOTS WOULD LIE IN LINE WITH EACH OTHER. | 114 |
| FIGURE 5-9. EXAMPLE OF A QQ PLOT WHERE 2D (X) VS. 3D (Y) MEASUREMENT RESULTS ARE PERFECTLY CORRELATED..... | 115 |
| FIGURE 5-10. HISTOGRAM OF SCENE A ROOF EDGE MEASUREMENTS FOR 2D PARTICIPANTS (TOP, N = 21) AND 3D PARTICIPANTS (BOTTOM, N = 22). BIN SIZE = 0.1M. | 117 |
| FIGURE 5-11. BOXPLOTS OF SCENE A MEASUREMENTS FOR 2D PARTICIPANTS (LEFT, N = 21, Y RANGE FROM 12.27M TO 13.18M) AND 3D PARTICIPANTS (RIGHT, N = 22, Y RANGE FROM 11.33M TO 13.17M)..... | 118 |
| FIGURE 5-12. THEORETICAL VS. SAMPLE QUANTILE PLOTS FOR SCENE A 2D (TOP) AND 3D (BOTTOM). X AXIS: 0 = MEAN (50% FREQUENCY) AND DEVIATIONS FROM THIS REPRESENT ONE QUANTILE (OR 25%). Y AXIS = MEASUREMENT RESULTS (IN M) MADE BY PARTICIPANTS. EACH POINT REPRESENTS AN OBSERVED MEASUREMENT MADE BY ONE PARTICIPANT. SOLID LINE DENOTES THE THEORETICAL NORMAL DISTRIBUTION. | 119 |
| FIGURE 5-13. 2D VS. 3D QUANTILE PLOTS FOR SCENE A, MEASUREMENT OF BUILDING ROOF EDGE. DASHED LINE REPRESENTS 1.0 AGREEMENT, WHERE 2D RESULTS= 3D RESULTS. | 121 |
| FIGURE 5-14. GENERAL PARTICIPANT RESPONSES TO QUESTION 8 – “HOW ACCURATELY DO YOU FEEL YOU MEASURED SCENE A?”. SCORES WERE RANKED ON A SCALE OF 1-7, WHERE 1 = NOT AT ALL ACCURATELY, 7 = VERY ACCURATELY. N.B. VALUES OF 0 [NO DATA INPUTTED] NOT INCLUDED ON SCALE. GRAPH SHOWS COMBINED ANSWERS FROM 2D AND 3D GROUPS, N = 46 (INCLUSIVE OF TWO PARTICIPANTS WHOSE QUANTITATIVE RESULTS WERE DISCOUNTED)..... | 122 |
| FIGURE 5-15. HISTOGRAMS OF THE SCENE B LONGEST CANOPY WIDTH MEASUREMENTS MADE BY 2D GROUP (TOP, N=22) AND 3D GROUP (BOTTOM, N=22). BIN SIZE = 1M..... | 125 |
| FIGURE 5-16. BOXPLOTS OF THE SCENE B LONGEST CANOPY WIDTH MEASUREMENTS MADE BY 2D GROUP (LEFT) AND 3D GROUP (RIGHT) | 126 |
| FIGURE 5-17. THEORETICAL VS. SAMPLE QUANTILE PLOTS FOR SCENE B 2D (TOP) AND 3D (BOTTOM). ALONG THE X AXIS, 0 = MEAN (50% FREQUENCY) AND DEVIATIONS FROM THIS ARE ONE QUANTILE (OR 25%). Y AXIS = MEASUREMENT RESULTS (IN M) MADE BY PARTICIPANTS. ONE POINT REPRESENTS A MEASUREMENT MADE BY ONE PARTICIPANT. SOLID LINES DENOTE THE THEORETICAL NORMAL DISTRIBUTION. | 128 |
| FIGURE 5-18. 2D VS. 3D QUANTILE PLOTS FOR SCENE B, ESTIMATED LONGEST CANOPY DIAMETER. DASHED LINE REPRESENTS 1.0 AGREEMENT. | 130 |

| | |
|---|-----|
| FIGURE 5-19. GENERAL PARTICIPANT RESPONSES TO QUESTION 8 – “HOW ACCURATELY DO YOU FEEL YOU MEASURED SCENE B?”. SCORES WERE RANKED ON A SCALE OF 1-7, WHERE 1 = NOT AT ALL ACCURATELY, 7 = VERY ACCURATELY. GRAPH SHOWS COMBINED ANSWERS FROM 2D AND 3D GROUPS, N = 46 (INCLUSIVE OF TWO PARTICIPANTS WHOSE QUANTITATIVE RESULTS WERE DISCOUNTED). | 131 |
| FIGURE 6-1. PLAN VIEW OF SCENE C POINT CLOUD, SHOWING THE LOCATION OF THE POINTS OF INTEREST THAT WERE INTERPRETED BY THE PARTICIPANT. POIS ARE HIGHLIGHTED BY FLASHING PINK GLOW. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013)..... | 145 |
| FIGURE 6-2. PLAN VIEW OF SCENE D POINT CLOUD, SHOWING THE LOCATION OF THE POINTS OF INTEREST THAT WERE INTERPRETED BY THE PARTICIPANT. POIS ARE HIGHLIGHTED BY FLASHING PINK GLOW. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013)..... | 145 |
| FIGURE 6-3. PHOTOGRAPHS SHOWING VIRTUAL REALITY THEATRE SET-UP, DURING INTERPRETATION TASK. PARTICIPANT POSED BY MODEL. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). | 155 |
| FIGURE 6-4. PLAN AND SIDE VIEWS OF SCENE C, WHICH MEASURES (WIDTH X LENGTH X HEIGHT) APPROX. 53M X 35M X 18M. THE POINTS CLOUD IS MADE UP OF 1563 GROUND POINTS, 2450 HIGH VEGETATION POINTS, AND 2991 BUILDING POINTS. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013B). | 156 |
| FIGURE 6-5. PLAN AND SIDE VIEWS OF SCENE D, WHICH MEASURES WIDTH X LENGTH X HEIGHT) APPROX. 50M X 35M X 75M. THE POINTS CLOUD IS MADE UP OF 6209 GROUND POINTS, 7242 VEGETATION, AND 1123 BUILDING POINTS. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A). | 157 |
| FIGURE 6-6. PLOTS COMPARING THE INTERPRETATION RESULTS (ACCURACY OUT OF 100%, Y AXES) FOR 2D VS. 3D FOR EACH POI (X AXES). SOLID RECTANGLE DENOTES SIGNIFICANT DIFFERENCE IN PROPORTIONALITY BETWEEN 2D AND 3D INTERPRETATION RESULTS. DASHED RECTANGLE DESCRIBES THE SAME, BUT AT A 90% CONFIDENCE LEVEL..... | 168 |
| FIGURE 6-7. GRAPHS SHOWING THE OVERALL LEVEL OF PARTICIPANT CONFIDENCE WHILE CARRYING OUT THE INTERPRETATION TASKS FOR SCENE C (LEFT) AND SCENE D (RIGHT). SCORES WERE GIVEN IN RESPONSE TO FEEDBACK Q10. RESULTS REPRESENT THE GENERAL COMBINED CONFIDENCE SCORES FOR BOTH 2D AND 3D METHODS, FOR EACH SCENE. X AXIS SCALE: 1 = VERY CONFIDENT..., 7 = NOT AT ALL CONFIDENT ABOUT THE ANSWERS GIVEN WHEN IDENTIFYING THE FEATURE OF THE FLASHING POINTS. Y AXIS: % OF PARTICIPANTS (TOTAL OF 46 FOR EACH SCENE). | 173 |
| FIGURE 6-8. INTERPRETATION TASK FINDINGS, SHOWING THE DIFFERENCE BETWEEN 2D AND 3D INTERPRETATIONS OF THE TEN POINTS OF INTEREST. THERE WERE NO FEATURES THAT HAD A BETTER 2D ACCURACY THAN 3D. | 178 |

List of Tables

| | |
|--|----|
| TABLE 0-1. GLOSSARY LIST. | XV |
| TABLE 1-2. KEY AIMS OF THE STUDY. | 7 |
| TABLE 2-1. INTERPRETATION TASKS SPECIFIC TO VISUAL IMAGE ANALYSIS, IDENTIFIED FROM CONTENT ANALYSIS OF 16 TEXTS (DATED 1922 – 1960) REGARDING HUMAN INTERPRETATION OF REMOTELY-SENSED AERIAL IMAGERY. TAKEN FROM BIANCHETTI AND MACEACHREN, 2015. | 18 |

| | |
|---|-----|
| TABLE 2-2. LIST OF GAPS FROM THE LITERATURE REVIEW, NUMBERED IN ORDER OF APPEARANCE. | 32 |
| TABLE 2-3. RESEARCH QUESTIONS AND HYPOTHESES..... | 34 |
| TABLE 3-1. SPECIFICATION OF AIRBORNE LASER-SCANNING (ALS) DATASETS ACQUIRED OVER BRISTOL AND LONDON STUDY SITES | 38 |
| TABLE 3-2. LIST OF OPEN-SOURCE AND COMMERCIAL SOFTWARE PACKAGES USED DURING DATA EDITING FOR AOIS AND VISUALISATION. | 41 |
| TABLE 3-3. GROUND POINTS CLASSIFICATION ROUTINE (MACRO STEP 1), WHICH CONDITIONALLY ALLOCATES THE LIDAR POINTS TO GROUND AND A (TEMPORARY) VEGETATION CLASS. A PSEUDOCODE EXPLANATION IS ALSO PROVIDED. NUMBERS IN BRACKETS REFER TO THE NON-STANDARD LIDAR CLASSIFICATION OF THE ALS SENSOR FORMAT. *FOR SCENE C, CLASSIFICATION IS ALSO RUN WITH SOURCE CLASSES SET AT SECOND AND THIRD. | 43 |
| TABLE 3-4. BUILDING CLASSIFICATION ROUTINE (MACRO STEP 2), WHICH FURTHER SORTS THE VEGETATION LIDAR POINTS FROM THE MACRO 1 OUTPUT INTO BUILDING AND VEGETATION CLASSES. A PSEUDOCODE EXPLANATION IS ALSO PROVIDED. NUMBERS IN BRACKETS REFER TO THE NON-STANDARD LIDAR CLASSIFICATION OF THE ALS SENSOR FORMAT. Z ACCURACY IS DEVIATION FROM THE ROOF PLANE. | 44 |
| TABLE 3-5. A MEASURE POINT DENSITY TOOL (TERRASOLID, 2013A) WAS USED TO ANALYSE THE DENSITY OF THE CLASSIFIED POINT CLOUDS. GREY CELLS INDICATE ABSENCE OF FEATURE FROM A SCENE. SCENE C VEGETATION WAS NOT MEASURED. | 46 |
| TABLE 3-6. DATA FILES USED DURING GENERATION OF SCENE A, B, C AND D DATASETS, USING TERRASCAN SOFTWARE (TERRASOLID, 2013A). THE .TXT FILES USED IN THE VISUALISATION SYSTEM INCLUDED GROUND, BUILDING, AND HIGH VEGETATION (OVER 2M). LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013A & 2013B)..... | 47 |
| TABLE 3-7. HARDWARE REQUIRED FOR VISUALISATION SET-UP | 50 |
| TABLE 3-8. SUMMARY OF BASIC BACKGROUND CHARACTERISTICS OF THE FOUR PILOT PARTICIPANTS. SOME CHARACTERISTICS, E.G. NATIONALITY, ARE OMITTED FROM THE TABLE EITHER BECAUSE OF DATA LOSS OR TO PRESERVE THE ANONYMITY OF THE VOLUNTEERS..... | 59 |
| TABLE 3-9. LIST OF DATA OUTPUTS FROM RELEVANT PHASES OF THE EXPERIMENT. UNDERLINED TEXT DENOTES DATA USED IN RESULTS. STEREO DATA ARE ONLY USED TO DETERMINE PARTICIPANT INCLUSION AND VISUALISATION SET-UP. | 73 |
| TABLE 3-10. LIST OF SOFTWARE USED FOR PROCESSING DATA GENERATED BY EXPERIMENT. | 75 |
| TABLE 4-1. RANKING OF PARTICIPANTS FOR LIDAR/LASER-SCANNING (A) KNOWLEDGE AND (B) EXPERIENCE, BASED ON TRANSCRIPTIONS. WHERE 0 = NOVICE, 1 = INFORMED, 2 = EXPERT..... | 81 |
| TABLE 4-2. SCALE OF SCORE CARD RESULTS FOR THE RANDOT™ TEST (STEREO OPTICAL CO., 2009). NUMBER AND PERCENTAGE OF PARTICIPANTS (MALE AND FEMALE). A SCORE OF 6/10 OR ABOVE LEAD TO THE VOLUNTEER'S INCLUSION IN THE FULL VISUALISATION EXPERIMENT. | 87 |
| TABLE 4-3. PARAMETRIC T-TEST BETWEEN FEMALES AND MALES FOR ALL STEREOACUITY SCORES AND THOSE OVER THAN OR EQUAL TO 6/10..... | 89 |
| TABLE 5-1. INSTRUCTIONS TOLD TO PARTICIPANTS DURING THE MEASUREMENT TASK. VERBAL AND VISUAL DIRECTIONS WERE GIVEN..... | 110 |
| TABLE 5-2. DISTRIBUTION OF MEASUREMENTS IN THE 2D AND 3D GROUPS FOR SCENE A'S ROOF EDGE..... | 116 |

| | |
|--|-----|
| TABLE 5-3. SHAPIRO-WILK NORMALITY TEST FOR 2D AND 3D GROUP, SCENE A. | 120 |
| TABLE 5-4. FURTHER COMMENTS ON FEEDBACK QUESTION 8, FROM PARTICIPANTS WHO MEASURED SCENE A IN 2D. ANY COMPARISONS ARE IN REFERENCE TO SCENE B IN 3D. ... | 123 |
| TABLE 5-5. FURTHER COMMENTS ON FEEDBACK QUESTION 8, FROM PARTICIPANTS WHO MEASURED SCENE A IN 3D. ANY COMPARISONS ARE IN REFERENCE TO SCENE B IN 2D. ... | 123 |
| TABLE 5-6. DISTRIBUTION OF MEASUREMENTS IN THE 2D AND 3D GROUPS FOR SCENE B'S CANOPY WIDTH..... | 126 |
| TABLE 5-7. SHAPIRO-WILK NORMALITY TEST FOR 2D AND 3D GROUP, SCENE B. | 129 |
| TABLE 5-8. SUMMARY OF QUANTITATIVE MEASUREMENT TASK RESULTS. | 132 |
| TABLE 5-9. SUMMARY OF QUALITATIVE MEASUREMENT TASK RESULTS. | 133 |
| TABLE 6-1. SUMMARY OF REFERENCE ANSWERS BASED ON LIDAR ACQUISITION INTERPRETATION BY ONE EXPERT OPERATOR. DATA ACQUISITION DATES ARE LISTED FOR LIDAR, ORTHOPHOTOS AND GOOGLE STREETVIEW IMAGERY..... | 147 |
| TABLE 6-2. FEATURE VERIFICATION OF POINT OF INTEREST (POI) A IN SCENE C. THE WHITE DOT (JUST SEEN) IN THE ORTHOPHOTO SHOWS THE POI LOCATION. THE $A_1 - A_2$ LINE IN THE ORTHO DATA IS SHOWN AS 2D TRANSECT CROSS-SECTION IN THE LIDAR DATA. THE POI IS LOCATED WITHIN THE WHITE CIRCLE. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTO © GEOPERSPECTIVES (2013). UN-SCALED SCREENSHOT IMAGES TAKEN FROM TERRASOLID LIDAR SOFTWARE GUI (TERRASOLID, 2013B)..... | 148 |
| TABLE 6-3. FEATURE VERIFICATION OF POINT OF INTEREST (POI) B IN SCENE C. THE WHITE DOT (JUST SEEN) IN THE ORTHOPHOTO SHOWS THE POI LOCATION. THE $B_1 - B_2$ LINE IN THE ORTHO DATA IS SHOWN AS 2D TRANSECT CROSS-SECTION IN THE LIDAR DATA. THE POI IS LOCATED WITHIN THE WHITE CIRCLE. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTO © GEOPERSPECTIVES (2013). UN-SCALED SCREENSHOT IMAGES TAKEN FROM TERRASOLID LIDAR SOFTWARE GUI (TERRASOLID, 2013B)..... | 149 |
| TABLE 6-4. FEATURE VERIFICATION OF POINT OF INTEREST (POI) C IN SCENE C. THE WHITE DOT (JUST SEEN) IN THE ORTHOPHOTO SHOWS THE POI LOCATION. THE $C_1 - C_2$ LINE IN THE ORTHO DATA IS SHOWN AS 2D TRANSECT CROSS-SECTION IN THE LIDAR DATA. THE POI IS LOCATED WITHIN THE WHITE CIRCLE. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTO © GEOPERSPECTIVES (2013). UN-SCALED SCREENSHOT IMAGES TAKEN FROM TERRASOLID LIDAR SOFTWARE GUI (TERRASOLID, 2013B)..... | 150 |
| TABLE 6-5. FEATURE VERIFICATION OF POINT OF INTEREST (POI) E IN SCENE C. THE WHITE DOT (JUST SEEN) IN THE ORTHOPHOTO SHOWS THE POI LOCATION. THE $E_1 - E_2$ LINE IN THE ORTHO DATA IS SHOWN AS 2D TRANSECT CROSS-SECTION IN THE LIDAR DATA. THE POI IS LOCATED WITHIN THE WHITE CIRCLE. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTO © GEOPERSPECTIVES (2013). UN-SCALED SCREENSHOT IMAGES TAKEN FROM TERRASOLID LIDAR SOFTWARE GUI (TERRASOLID, 2013B)..... | 151 |
| TABLE 6-6. FEATURE VERIFICATION OF POINT OF INTEREST (POI) G IN SCENE D. THE WHITE DOT (JUST SEEN) IN THE ORTHOPHOTO SHOWS THE POI LOCATION. THE $G_1 - G_2$ LINE IN THE ORTHO DATA IS SHOWN AS 2D TRANSECT CROSS-SECTION IN THE LIDAR DATA. THE POI IS LOCATED WITHIN THE WHITE CIRCLE. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTO © GEOPERSPECTIVES (2013). UN-SCALED SCREENSHOT IMAGES TAKEN FROM TERRASOLID LIDAR SOFTWARE GUI (TERRASOLID, 2013B)..... | 152 |
| TABLE 6-7. FEATURE VERIFICATION OF POINT OF INTERESTS (POI) I AND J IN SCENE D. THE WHITE DOT (JUST SEEN) IN THE ORTHOPHOTO SHOWS THE POI LOCATION. THE POIS IS LOCATED WITHIN THE WHITE CIRCLE. LIDAR DATA © AIRBUS DEFENCE AND SPACE LTD. (2013). ORTHOPHOTO © GEOPERSPECTIVES (2013). UN-SCALED SCREENSHOT IMAGES TAKEN FROM TERRASOLID LIDAR SOFTWARE GUI (TERRASOLID, 2013B)..... | 153 |

| | |
|---|-----|
| TABLE 6-8. QUESTION 10, WHICH WAS POSED TO PARTICIPANTS AFTER COMPLETION OF THE INTERPRETATION TASK..... | 158 |
| TABLE 6-9. EXAMPLES OF THE ACCEPTED AND REJECTED INTERPRETATION ANSWERS FOR EACH POINT OF INTEREST (POI), FROM THE SCENE C (POIs A-E) AND SCENE D (POIs F-J)..... | 160 |
| TABLE 6-10. NUMBER OF PARTICIPANTS TAKING PART IN THE INTERPRETATION TASK..... | 165 |
| TABLE 6-11. P-VALUE RESULTS OF 2D VS. 3D INTERPRETATIONS FOR EACH POI (A-J), USING A TWO-TAILED 2-SAMPLE TEST FOR EQUALITY OF PROPORTIONS WITH CONTINUITY CORRECTION (PROP.TEST() FUNCTION IN R (R CORE TEAM, 2014). BOLD P-VALUES DENOTE A SIGNIFICANT RESULT BETWEEN THE 2D AND 3D GROUPS, AT 95% CONFIDENCE LEVEL, I.E. <0.05. | 167 |
| TABLE 6-12. BINOMIAL DISTRIBUTION RESULTS FOR 2D AND 3D INTERPRETATION OF POIs A TO J. WHERE X IS THE NUMBER OF CORRECT INTERPRETATIONS, N IS THE TOTAL NUMBER OF OBSERVATIONS, AND P IS THE PROBABILITY OF SUCCESS (X/P). P-VALUE IS GIVEN FOR THE RESULTS OF THE “SUCCESS IS GREATER”-TAILED BINOM.TEST() PERFORMED IN R, AT THE 95% CONFIDENCE LEVEL. POIs WITH P-VALUES BELOW 0.05 DISPLAY BIAS TOWARDS BEING CORRECT AND ARE SHOWN IN BOLD. NORMAL DISTRIBUTION FOR THE GIVEN N IS ALSO SHOWN..... | 170 |
| TABLE 6-13. SUMMARY OF MEAN INTERPRETATION ACCURACY PER AOI, BASED ON MEAN POI ACCURACY..... | 173 |
| TABLE 6-14. SUMMARY OF THE RELATIONSHIP BETWEEN 2D VS. 3D INTERPRETATION ACCURACY FOR TEN POI FEATURES AND PARTICIPANT FEEDBACK REGARDING EACH SCENE. THERE WERE NO CASES WHERE 2D > 3D..... | 176 |

List of Equations

| | |
|------------------|-----|
| EQUATION 1..... | 12 |
| EQUATION 2 | 108 |
| EQUATION 3..... | 161 |
| EQUATION 4. | 163 |

Glossary

A list of acronyms and terms that are used within the text are defined in Table 0-1. Note also the convention of terms, regarding 2D and 3D, in section 1.2, ahead of the literature review.

Table 0-1. Glossary list.

| Acronym / term | Definition |
|--------------------------|---|
| 2D | Two-dimensional. In this study, same definition as 2.5D. |
| 2.5D | Two-dimensional representation of a 3D object, using depth cues |
| 3D | (Stereoscopic) three-dimensional |
| AOI | Area of interest |
| ALS | Airborne laser scanning |
| binocular / binocular | Relating to two eyes |
| CAD | Computer-aided design |
| CPU | Central Processing Unit – main computer component that carries out programs |
| DOF | Degrees of freedom |
| FOV | Field of view or field of vision |
| FPS | Frames per second |
| GI | Geographic information |
| GIS | Geographical Information System or Geographical Information Science |
| GUI | Graphical user interface |
| HCI | Human-computer interaction |
| HMD | Head-mounted display |
| IPD | Interpupillary distance – distance between the eyes |

| | |
|-----------|-----------------------------|
| IQR | Inter-quartile range |
| lidar | Light detection and ranging |
| monocular | Relating to one eye |
| POI | Point of Interest |
| QA | Quality assurance (of data) |
| QC | Quality control |
| UAV | Unmanned aerial vehicle |
| TLS | Terrestrial laser scanning |
| VE | Virtual environment |
| VR | Virtual reality |

1.

Introduction

1. Introduction

Day-to-day, we view our 3D environment with the stereoscopic depth that is gained from the overlap of two two-dimensional images – one from each eye. When viewing information that is displayed via a single screen, the representation is comparatively flat and two-dimensional. Although stereoscopic displays exist, the concept of immersive 3D data visualisation has not been adopted by the remote-sensing community. This study therefore questions whether stereo 3D, as opposed to 2D, visualisation can add value to the analysis of remotely-sensed data.

As humans, we have the ability to see with stereoscopic depth because our eyes are set slightly apart. This allows us to view objects from two slightly offset angles, Figure 1-1. Each eye has a field of view (FOV), or field of vision, which is considered to have an extent of 180° horizontally and 140° vertically (Gibson, 1979). The shaded area of Figure 1-1 highlights the intersection of the left and right FOVs, where binocular vision happens. Where this overlap occurs, the brain integrates the image from each eye and perceives them as three-dimensional objects (Howard and Rogers, 2012).

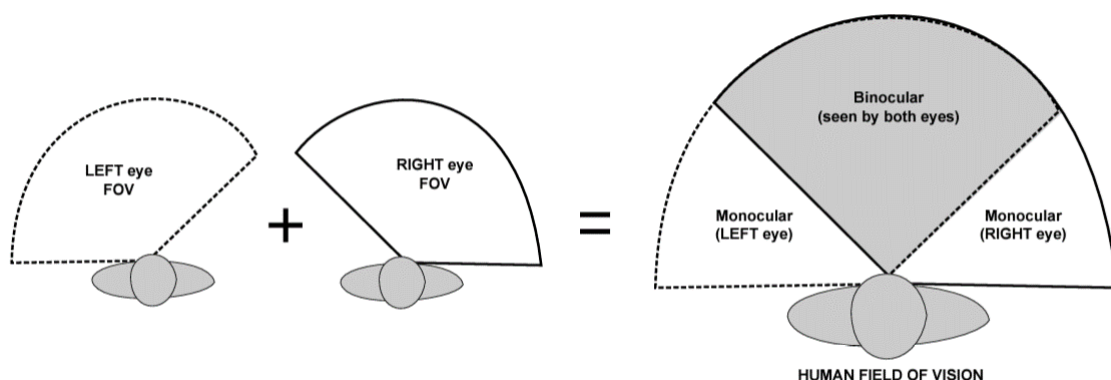


Figure 1-1. Illustration of the human field of vision, shown from plan view.

Certain technologies have been developed to trick the brain into 'seeing' representations displayed on flat surfaces as being in 3D. The 3D displays take advantage of human vision and from the disparity between the left and right eye, the brain gains binocular cues. An example of 2D vs. 3D stereoscopic

visualisation is shown in Figure 1-2. From the 2D images, the viewer might use relative size difference as a cue to infer the distance between the objects and themselves. For example, one might assume that the smaller object is further away from the viewer than the larger object. In contrast, when perceived using binocular vision, one display per eye, the object can be perceived in their true 3D form and positions, which are behind and in front of the display.

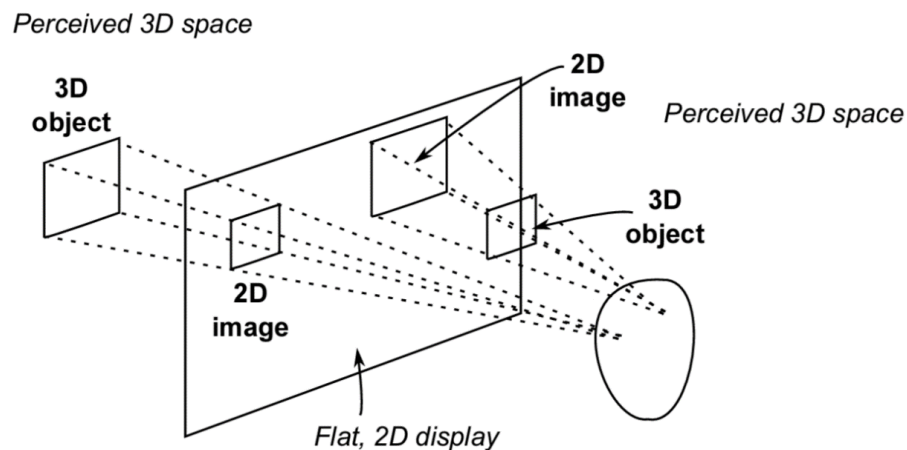


Figure 1-2. The visual perception of objects in 3D space, compared to their pictorial representation on a flat screen. Oval represents human head. Adapted from Hodges (1992, cited in Howard and Rogers, 2012, p.539).

It is important to acknowledge at this stage that depth perception can, to some degree, be gained from monocular cues, such as relative sizing and occlusion of objects. These are pictorial cues that can be appreciated with one eye and these visual prompts are present when the view is static (and moving). A painter might use these to give the impression of depth in their artwork. An overview of visual cues are summarised in Figure 1-3, alongside their availability under 2D/3D visualisation conditions. Depth information is also provided by responses of the eye (its focus, rotation, and pupil size) in reaction to the stimulus (Tovée, 1996), and further detail on oculomotor and visual cues are described in Carr and England (1995) and Eysenck (2001). This study focuses on the added depth provided by stereopsis, which is only available during 3D visualisation and can enhance our understanding of the spatial arrangement of the displayed information. A 2D display will always lack this depth cue, regardless of any other cues present.

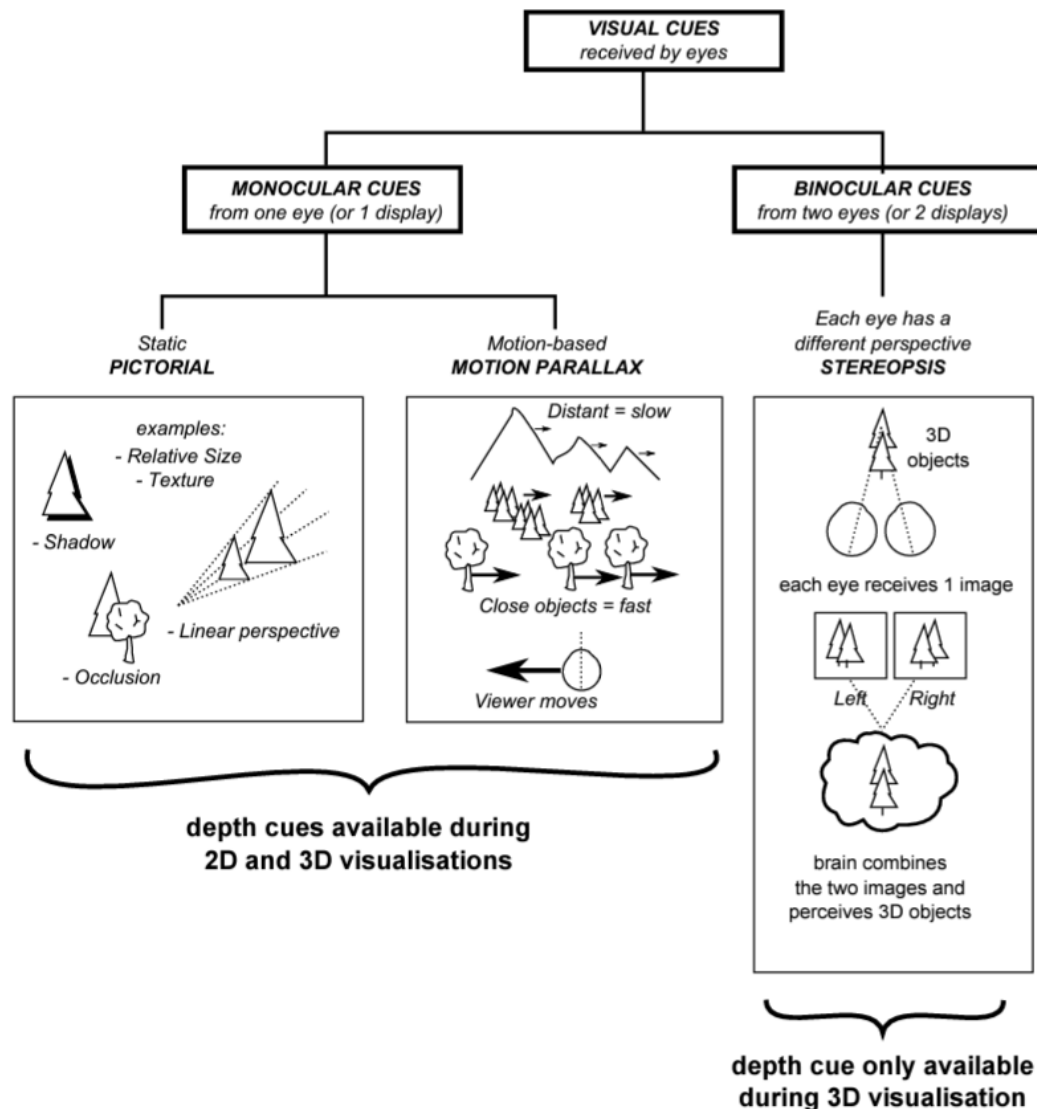


Figure 1-3. Summary of visual depth cues. Additional pictorial cues are described in Tovée (1996).

Increasingly, the visual information that we consume is represented through electronic devices, such as television, computer screens, and mobile devices. The use of monocular pictorial visual cues on these flat digital screens aids human spatial understanding of the information that is represented. The process of digital information visualisation is illustrated in Figure 1-4, adapted from Ware (2012), whereby the human viewing the information may be undertaking data manipulation and exploration. However, visual and cognitive processing, and subsequent manual data tasks, may be affected if the data were represented differently, via a binocular immersive display.

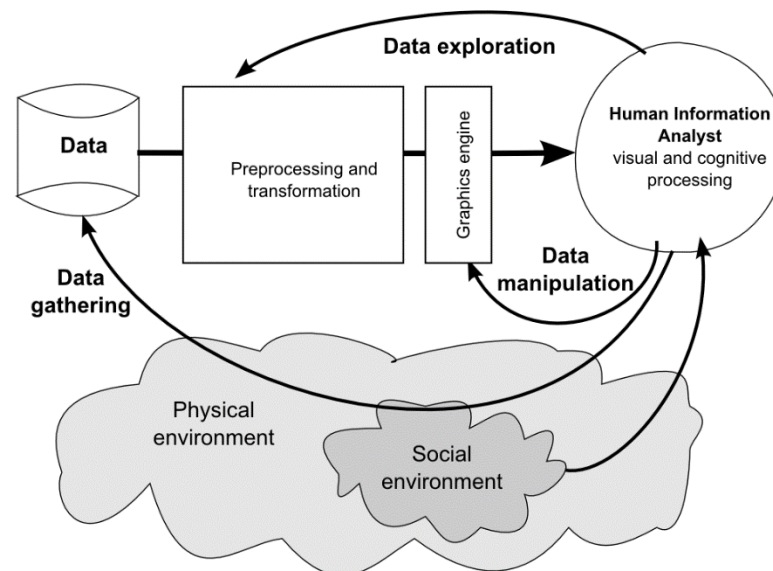


Figure 1-4. Schema showing the process of Visualisation (adapted from Ware, 2012)

Stereoscopic 3D representation, which uses both of the viewer's monocular FOVs (and combined binocular FOV), gives an immersive 3D view of a dataset. In effect, a virtual world is perceived and viewed, which transforms user observation into user experience, with the feeling of being within that environment (Holtzman, 1994). The concept of virtual reality (VR) is considered to have originally involved full immersion of the user in the hardware (Lin and Batty, 2009), as was first introduced with Ivor Sutherland's 1960s head-mounted display (Sutherland, 1968). In the 1980s and 1990s, as practical hardware advanced (Vince, 1998), 'virtual reality' became a buzzword (Faust, 1995). Although stereoscopic immersive visualisation was theoretically robust, the technology was still unable to allow comfortable cognitive translation of images into the perception of 3D objects. A lull in the interest of VR followed, but stereo technology has since made an impact via 3D movies in cinemas and 3DTVs in the home. The dwindling usage of the latter by the general public saw the decommissioning of 3D-dedicated channels from the UK entertainment television channels BBC (BBC, 2013) and Sky (Sweney, 2015). There have, however, been technological advancements in the computer gaming industry (Edge® 2013), such as development of the wearable Oculus VR headset (Oculus VR, 2015a; Oculus VR, 2015b), which allow a more comfortable immersive stereo visualisation technique. Despite the resurgence of interest in

3D visualisation technology by developers, there is a risk that its application is not targeted towards the visual information that demands or requires 3D representation.

Unlike the passive 3D viewing of entertainment content or the indiscriminate application of 3D techniques, the scientific community uses immersive stereo imaging as a means to actively answer questions or diagnose problems concerning 3D data. Stereoscopic human-data interaction is well considered and established in the fields of medicine (Walter *et al.*, 2010; Thomsen *et al.*, 2005; McAuliffe *et al.*, 2001). According to Walter *et al.* (2010), the application of new visualisation techniques in the medical domain reportedly permits a more adequate representation of massive datasets, which are not sufficiently offered by 2D images or software. The geosciences also exploit three-dimensional visualisation to understand geological and geomorphological structures (McCaffrey *et al.*, 2005; Trinks *et al.*, 2005; Jones *et al.*, 2008; Jones *et al.*, 2009; Bernardin *et al.*, 2011). Despite the suitability of three-dimensional geographical data for stereoscopic display, namely laser-scanned point clouds (example shown in Figure 1-5), the field of remote-sensing has been slow to adopt the immersive technique to facilitate the exploration, measurement, and interpretation of these datasets.

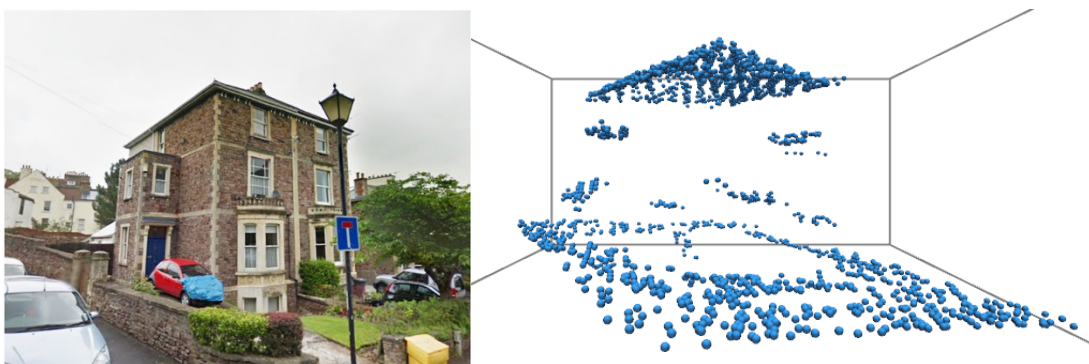


Figure 1-5. Ground truth image (Google, 2014), left, depicting suburban feature and, right, its equivalent laser-scanned point cloud. Lidar data © Airbus Defence and Space Ltd. (2013a). Data were captured using an Optech Gemini airborne sensor and included overlapping flightlines (see Appendix A) with ~ 8 points per metre squared (ppm^2). Right-hand image shows processed data, with building points at a density of $\sim 7 \text{ppm}^2$ and ground points at $\sim 2 \text{ppm}^2$. The full lidar processing method is detailed in Chapter 3.

However, there is some evidence that suggests 3D visualisation of scanned point clouds should be exploited by the remote-sensing community. The few virtual lidar point cloud studies have applied the visualisation method to forested and geomorphological environments (Warner *et al.*, 2003; Kreylos *et al.*, 2006; Kreylos *et al.*, 2008; Burwell *et al.*, 2012), and Yan *et al.* (2015) highlight its potential application to urban land cover classification. As illustrated by the images in Figure 1-5, the laser-scans are a 3D representation of the environment. Given the success of 3D data visualisation in other areas of visual scientific analysis, the representation of three-dimensional point clouds within a 3D environment would be logical and, if it significantly adds value to 2D analysis, its adoption justified.

1.1 Research aim

The study seeks to assess whether an immersive stereoscopic 3D gives added-value to the visual analysis of three-dimensional lidar data, compared to a 2D display. The general aim of the research project is broken down into three sections, see Table 1-2, which cover technical development, the execution of the participant experiment, and the evaluation of the results.

Table 1-2. Key aims of the study.

| Aim | Description |
|------------|---|
| TECHNICAL | Develop a 3D visualisation system that allows users to carry out manual lidar point cloud tasks for different geographical features and scenes. Design an experiment around research questions. |
| EXPERIMENT | Test Human performance during 3D and 2D visualisation tasks that simulate lidar tasks |
| EVALUATION | Determine the accuracy and/or precision of 3D task results against 2D task results. Reflect on the experiment design. |

1.2 Thesis structure

The thesis structure is presented in Figure 1-6. The introduction, Chapter 1, gives an overview of 3D human vision and lidar data. Chapter 2, the literature review, describes the visualisation of lidar point clouds - its importance, the standard 2D technique used, and considers the proposed 3D technique. Gaps in current knowledge are flagged up throughout the text and these help the formation of the research questions.

Novel methodological approaches were developed for the study and these are detailed in Chapter 3. A clear workflow is presented, beginning with experiment development, the pilot study, data collection, data processing, and analysis.

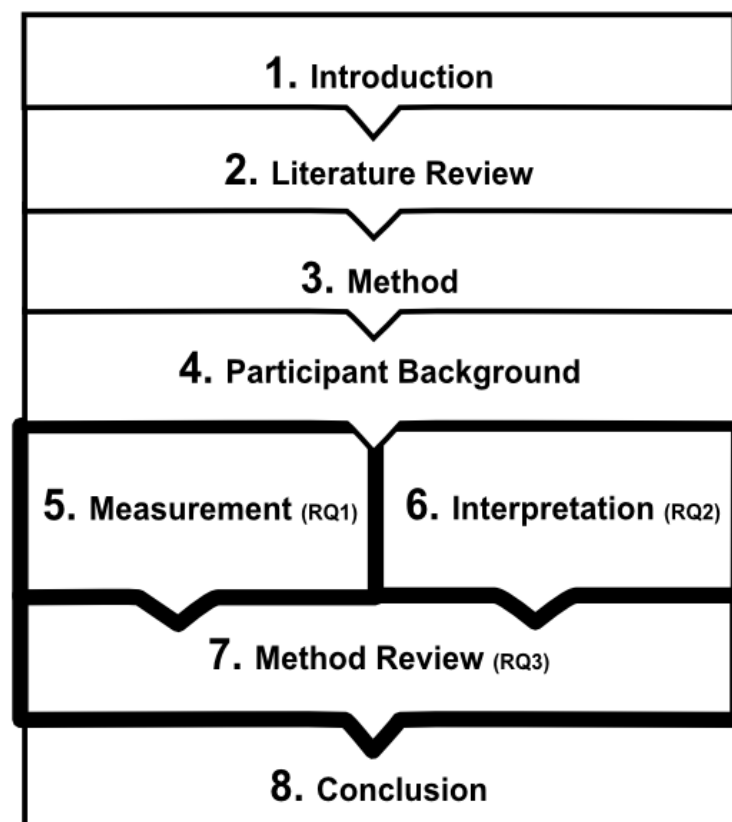


Figure 1-6. Thesis Structure. Bold elements denote research questions (RQ).

Participant background characteristics are presented in Chapter 4, to give an overview of the sample of volunteers. Chapter 5, which describes the

measurement task, outlines further methods specific to that task, followed by a synopsis of the results. These are further discussed in relation to the wider literature and the methodological approach is reviewed. The interpretation task chapter, Chapter 6, is presented similarly to the measurement chapter.

The method review, Chapter 7, considers the methodological approach of the general method. Recommendations are suggested from reflections of the experimental approach, inclusive of those derived from the two tasks. Chapter 8 brings together the main findings of the research, stating its impact and a conclusion to the study.

Throughout the study, unless otherwise stated, the term '3D' is used in this thesis to define 'three-dimensional' or 'third-dimension', in the stereoscopic sense. In other words, 3D is used to describe objects with xyz coordinates are presented in a 3D space. This is not to be confused with 2.5D, which describes the projection of a 3D object (or data) onto a flat, 2D plane, e.g. computer screen or paper. The representation of 3D objects through the medium of 2D, *i.e.* via a computer screen or on paper, is described hereafter as 2D.

2.

Literature review

2. Literature review

2.1 Introduction

This study is positioned between different fields of research, notably remote-sensing, geovisualisation, and cognitive psychology. The research is guided by literature that spans these disciplines; the literature review chapter critiques work of relevance and identifies gaps that are not currently addressed by any other published work. Section 2.4, at the end of the chapter, presents the study's research questions (RQs), which were generated as a direct result of the gaps found during the literature review.

2.2 Lidar point cloud visualisation

2.2.1 Background to lidar

The acronym, lidar, stands for light detection and ranging, which describes the use of light to actively measure the distance between the scanning instrument and the target object. The data are acquired using a laser instrument to scan the environment and these are processed to display a 3D representation of the scanned scene. Lidar, therefore, is an example of three-dimensional data that could benefit from three-dimensional visualisation.

An example of the data collection technique is shown in Figure 2-1 and its associated calculation is explained in Equation 1. This data acquisition method can be applied at different scales within a geographical context and Figure 2-2 summarises the current acquisition methods, including spaceborne (Zwally *et al.*, 2002; Gong *et al.*, 2011), airborne (Holmgren & Persson, 2004; Omasa *et al.*, 2008), mobile (Puttonen *et al.*, 2011; Sanz-Cortiella, 2011) and terrestrial (Buckley *et al.*, 2008; Park *et al.*, 2007). Alongside the increasing portability and compactness of laser scanners, unmanned aerial vehicles (UAVs) have been the more recent platform adopted for remotely-sensed data acquisition

(Anderson and Gaston, 2013), further increasing the accessibility of remote areas of interest.

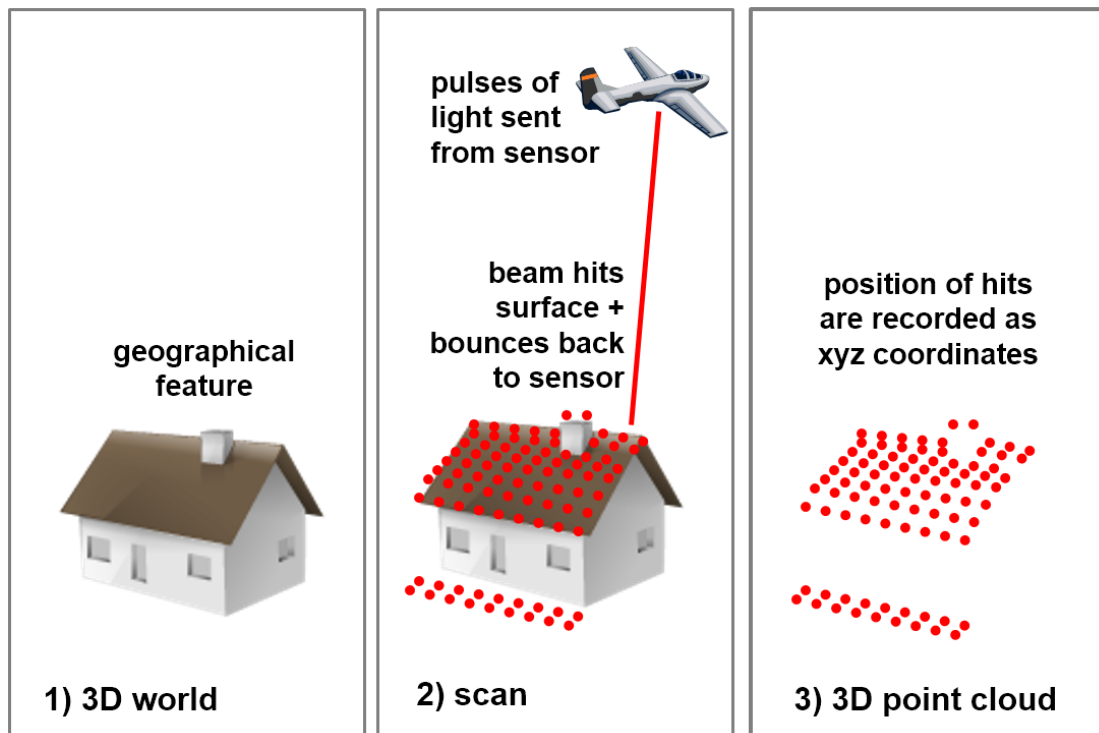


Figure 2-1. Image to show data acquisition example using laser-scanner mounted in a plane, scanning an urban geographical feature within the 3D physical environment.

$$\text{Distance from sensor} = (\text{Speed of light} / \text{Time taken for beam of light to return to sensor after hitting an object})/2$$

Equation 1.

The laser-scans can be acquired as discrete xyz points, known as a point cloud. Alternatively, scans can be captured using a full-waveform beam, resulting in a waveform return, which represents the strength of reflectance within each laser pulse (Söderman *et al.*, 2005). This research focuses on the point cloud, whose individual points, fundamentally, only describe the xyz position (and intensity) of each of the lidar returns. It is acknowledged that its raw, uninterpolated point cloud provides a powerful tool for understanding and

deriving parameters from the scanned structures, such as forest height and allometric estimations of carbon stock (Omasa *et al.*, 2007) or positional accuracy. It should be noted that the fidelity of the point cloud representation to the real-world environment can vary because of the intermittent point density of a laser-scanner. For example, Zimble *et al.* 2003 explains, using airborne laser-scanning of vegetation structure, that there is uncertainty in point cloud data because of the voids of data in the gaps between the points. This means that a raw point cloud and data derived from it, cannot truly be representative of the real-world environment. However, if the resolution or density of the points is high, then the detail of the scanned surface will be represented in finer detail. In section 3.2.1 (lidar data preparation) and 5.2.1 (part of the measurement task method), data density is discussed in relation to the data used in this study.

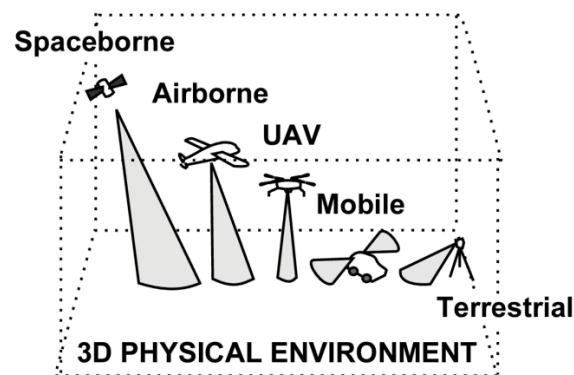


Figure 2-2. Illustration of laser-scanning data acquisition techniques. Image not to (relative) scale. From left to right, laser sensor mounted on satellite, aircraft/helicopter, drone (Unmanned Aerial Vehicle), road vehicle, tripod. Data capture methods vary between instruments.

As an alternative to lidar point clouds for 3D structural information, Leberl *et al.* (2010) outline the merits of photogrammetry, which involves the use of two stereo images to generate a true 3D representation. The image-based approach is favoured over lidar point cloud data by the authors because it is a continuous data coverage, unlike lidar (Zimble *et al.*, 2003), and, at least from an aerial acquisition perspective, has a single workflow. However, Leberl *et al.* (2010) acknowledge that the advantages of lidar over a photogrammetric approach include the penetration of leaves and detection of wires, from ALS

surveys. Furthermore, a photogrammetry approach only uses vertical, top-down stereoscopic visual analysis of the lidar-derived raster data; the viewer is not able to change the viewpoint. Structure-from-Motion (SfM) photogrammetry is a low-cost remote-sensing method (Westoby *et al.*, 2012) that uses multiple overlapping spectral images to generate a 3D point cloud, from different viewpoints. These can be acquired using UAVs (Kato *et al.*, 2015) or any other offset movement between sequential image capture, offering an inexpensive technique for point cloud generation at different scales, which could be viewed in 2D or 3D. The structure of the SfM point cloud is similar to that of a lidar point cloud, but its 3D coordinates are derived from overlapping 2D images, instead of a xyz laser hit.

A raw lidar point cloud is represented in the left image of Figure 2-3, which illustrates how segmentation of the points into different categories creates thematic variables. The interpretation from raw points into more easily usable data can be achieved automatically (Meng *et al.*, 2010; Flood, 2001), using algorithms, and interactively, with manual interpretation. Both techniques group points into the same classification e.g. vegetation, ground, *etc.* Once the lidar data have been processed, parameters can be derived from the classified points, with which further analysis can be carried out, e.g. digital terrain modelling, vegetation analysis, and urban modelling.

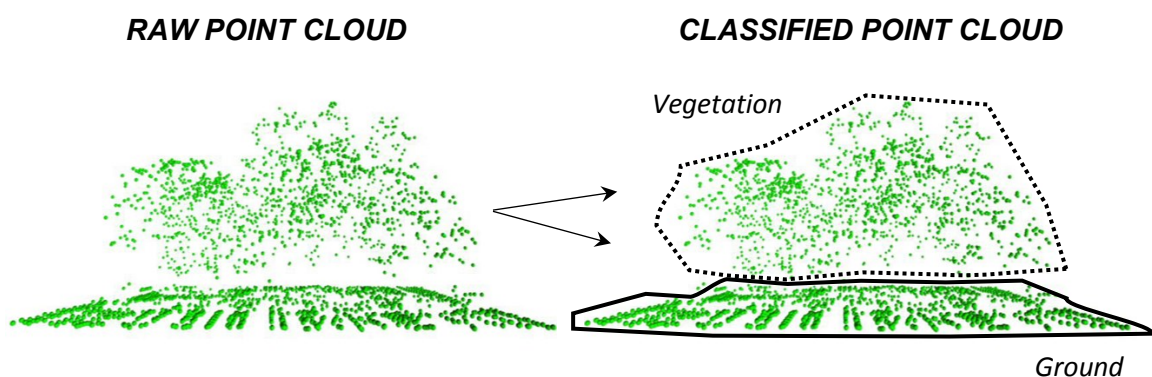


Figure 2-3. Illustration depicting a volume of raw point cloud (left), which is made up of xyz points collected via airborne laser-scanning. The raw data points can be segmented into different classifications (left), e.g. ground surface and vegetation. Diagram based on lidar data © Airbus Defence and Space Ltd. (2013a). Displayed density is ~ 2 points per metre squared (ppm^2) for ground points and $\sim 4ppm^2$ for vegetation.

2.2.2 Importance of point cloud visualisation

Discrete lidar datasets are made up of clouds of xyz points, as shown in Figure 2-3, and they can be manipulated automatically and manually. The automated and manual methods can complement one another (Gigli and Casagli, 2011), but analysis via automated systems, which involve the computer processing algorithms, is considered to be efficient (Haala and Kada, 2010) and reliable, with reproducible methods (Lovell *et al.*, 2011). Brodu and Lague (2012) state that automation is required in the processing of lidar data for their geomorphological study because of the size of the dataset and regard manual interaction as 'cumbersome' (Brodu and Lague, 2012, p.126). In response to the increase in the amount of data recorded during acquisition (Suomalainen *et al.*, 2011), the lidar community seeks to increase automated segmentation and interpretation. However, fully-automated processing and analysis is not always effective (Haala and Kada, 2010) and there continues to be a dependency on manual analysis, which is considered to be time-consuming and expensive, and varies between operators (Meng *et al.* 2010, p.839). The subjective nature of manual analysis means that only those features that appear to be significant are selectively investigated (Gigli and Casagli, 2011). Despite these shortcomings, 60-80% of commercial lidar production, shown in Figure 2-4 relies on manual classification and final quality checks (Flood, 2001). The schema could be interpreted in two ways: (a) manual interaction is important because it is relied on heavily during lidar production chain, commercial or otherwise, and (b) the automatic classification must be improved to reduce reliance on manual reclassification. Generally, the literature and industry fixate on (b) with the aim of reducing time and money spent on human operators. However, while repeating mantra (b) and attempting to reduce Flood's 2001 statistic, it is often overlooked that humans are and will always be a critical part of the process. Total auto-processing is not possible; a human is needed to oversee the workflow, based on the application, and set the parameters. Furthermore, the accuracy of automated algorithms falls short due to the heterogeneity of natural environments (Meng *et al.*, 2010). For example, Sithole and Vosselman (2005) found that their proposed segment-based automatic algorithm, as well as three other filters tested, returned significant

misclassification error when attempting to classify bare Earth points from a heavily vegetated, sloped study area (Sithole and Vosselman, 2005).

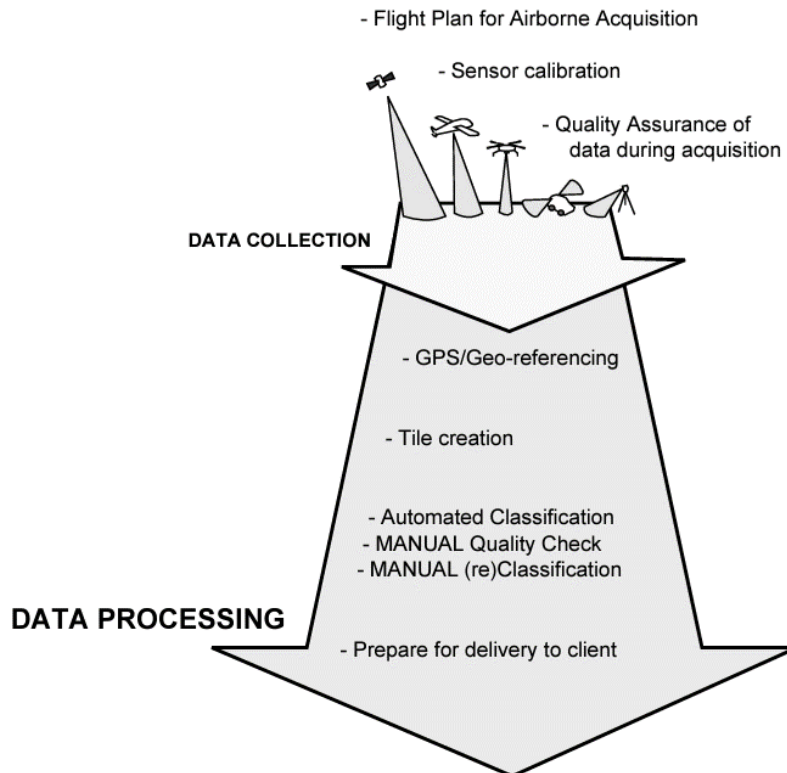


Figure 2-4. Flow diagram adapted from Flood (2001) showing the actions required during a commercial lidar data workflow, before delivery of the dataset to the client. Asterisks highlight the human aspects of the process, during which 60-80% of production occurs.

When automated processes fail, a manual approach is used to ‘tidy up’ the errors. In practice, the (semi-)automated systems require human-in-the-loop input before and after processing (Baltsavias, 2004; Helmut, 2008). During semi-automatic processing, operators would rather calibrate an algorithm that leaves unwanted points (Raber *et al.*, 2002) and this highlights the uncertainty of automated processing techniques. Furthermore, manual analysis of point clouds is commonly used when ground truth imagery is not available (Meng *et al.*, 2010). Regardless of the application of laser-scanning, there will always be a need for a human to check the quality of the data, depending on the requirements of the end-user. This is particularly relevant for large-scale heterogeneous environments. This research is underpinned by visual analysis approaches, which can enhance automated processes.

Visual tasks in lidar analysis

The laser-scanning literature alludes to certain analysis and visualisation tasks that may benefit from the use of stereoscopic immersive 3D environment. Alongside improving automated processing, Flood (2001) stated that better tools were needed during the visual quality checks (QC, or QA – ‘quality assurance’) of the data process workflow. Errors in the lidar data, such a disruption in the line of a building, can have a significant impact on the realism of the scanned scene in VR city modelling (Haala and Kada, 2010). Manual editing of façades is required for complex structures and detailed areas (Haala and Kada, 2010), especially in architectural style varies. This requires visual interpretation tasks, which (Bianchetti and MacEachren, 2015) describe in the context of aerial imagery interpretation, Table 2-1, overleaf, and can be applied to general cognitive remote-sensing data interpretation.

Lidar QA requires identification of misclassifications caused by the automated algorithms, followed by a judgement, *i.e.* ‘determining a characteristic of an image feature’ (Bianchetti and MacEachren, 2015). This manual QC task is commonly followed by reclassification of the point cloud, as shown in Figure 2-4, whereby the human operator will interactively correct the misclassified points. The different types of analyses that occur after segmentation (and QA) of the point cloud include manual extraction of parameters, which may include measurement of the relative sizes. In standard lidar visualisation software, there is commonly an interactive tool for measuring between points. In the GI software ArcGIS 10.1, it is suggested that three-dimensional distances are used to measure between trees and power lines (ESRI®, 2013), for utilities risk assessment applications. In engineering, manual interaction with the point cloud assists with the monitoring of infrastructure, such as pipeline displacement (Park *et al.*, 2007). For vegetated areas of interest, individual tree crown and stem parameters can be derived; the open-source lidar vegetation analysis software FUSION/LDV (McGaughey, 2014) allows the user to interactively derive structural metric information, which can contribute to canopy height modelling (Van Leeuwen *et al.*, 2011; Yang *et al.*, 2013).

Table 2-1. Interpretation tasks specific to visual image analysis, identified from content analysis of 16 texts (dated 1922 – 1960) regarding human interpretation of remotely-sensed aerial imagery. Taken from Bianchetti and MacEachren, 2015.

| Interpretation Task | Description (The process of...) |
|-----------------------|---|
| Search | visually scanning an image. |
| Detection | noticing an image feature. |
| Identification | recognizing an image feature. |
| Comparison | comparing two sources of information (image features, multiple images, or other types). |
| Judgment | determining a characteristic of an image feature. |
| Measurement | measuring the relative size of an image feature. |
| Signification | judging the importance or utility of an image feature to solving an analytical problem. |

Brodu and Lague (2012) demonstrated that, in their study of the classification of complex natural environments interactive, manual inputs could be used to train semi-automated processing of lidar point clouds. The point cloud also acts as the scaffolding of 3D models (Omasa *et al.*, 2008; Raber *et al.*, 2002) and the manual digitising required to achieve this (of features such as rooflines and roadside kerbs) requires the user to carry out similar cognitive interpretation tasks to linear measurement, *i.e.* determine and select the start and end of a line. This interpretation sub-task of 'signification' is used to judge the significance of a feature for problem-solving (Bianchetti and MacEachren, 2015) and it is also used during point cloud interpretation, when the user must determine the allocation of a point to a certain feature and, ultimately, classification. Of these manual lidar tasks, those that could benefit from the added depth perception of 3D display and those that are sufficiently executed in 2D are not determined by the literature.

Gap 1: Do not know which manual lidar point cloud analysis task(s) perform better in 3D, compared to 2D.

Perception of different variables

Environmental features can be crudely divided into two types - natural and man-made. Generally, natural features are not uniformly shaped and, unless dense, leaf-on vegetation is volumetric. In contrast, man-made features, like buildings, are commonly angular and planar. The fundamental geometrical differences between volumetric and planar features may affect the way that we can perceive them: consider a flat sheet of metal, which is a plane of two (xy) dimensions. If the sheet is bent and warped, more axes are required to describe the geometry of the data (Jones *et al.*, 2008). Jones *et al.* (2008) consider the terminologies used to describe the dimensionality of digital geoscientific datasets. They highlight that the more irregular a laser-scanned shape is, the more three-dimensional it becomes (even when data are displayed onto a 2D plane). This suggests that as the scanned feature becomes more irregular, and less planar, its dimensionality and the volumetric arrangement of the xyz points makes it more suitable for 3D visualisation.

In a study in which student participants were taught 3D skeletal hand anatomy through 2.5D displays (with dynamic vs. restricted views), Garg *et al.* (2002) highlight that its results are constrained by the planar stimulus. The authors imply that similar research into the visualisation of 3D objects should consider volumetric material, in order to widen the impact of results. This data dimensionality could translate through to larger-scale structures and, elsewhere, in the geographical domain, Seipel (2012) studied the effect of 2D, 'weak 3D' (2.5D), and 'strong 3D' (stereoscopic 3D) on human visualisation of 2D maps. However, the authors admit that the 18 participants viewed two-dimensional objects in a 3D space and therefore did not take advantage of the full stereoscopy by using 2D planar images. This leads to the assumption that planar buildings and continuous ground surfaces will not be represented as well in 3D as full volumetric scenes. Potentially, depending on the morphology of the data, certain features could be better represented than others in a stereoscopic projection. There is no existing research that addresses this hypothesis in a remote-sensing context. Furthermore, after Haklay (2002), are 3D representations of various features more favourable for visual lidar analysis than 2D display?

Gap 2: *Which lidar variables are better represented in 3D, as opposed to 2D?*

If a ground-filtering algorithm is automatically run on flat urban terrain and hilly vegetated terrain, depending on the parameters, there will be different accuracy outputs. In some cases of the latter, the vegetation may occlude underlying terrain. The algorithm must be manually tweaked to take into account the different environmental settings. Meng *et al.* (2010) explain that ground-filtering algorithms are site- or terrain-specific and it is not clear whether automated functions are appropriate in other areas. The perception of point clouds by the human cognitive system may also be affected by differences in terrain. Furthermore, it is not clear how the added depth of a 3D point clouds would affect human ability to process information from different environmental settings.

Gap 3: *The geometry of the data acquisition areas may affect visual 2D and 3D lidar visualisation analysis outputs.*

2.2.3 Standard 2D visualisation

Although there is an increasing availability of lidar data across fields, the current display of lidar is largely limited to viewing on a desktop computer screen. The 2D desktop environment requires an abstraction of the data (Neves *et al.*, 1997), whereby the representation of points projected onto a 2D screen. Figure 2-5 summarises how positional information from the 3D world is recorded as 3D data, but once it is visualised via a 2D display, such as a computer screen, information is lost before it can be exploited by the user. This is the current accepted method for viewing, interrogating, editing, and analysing datasets. The rendering of 3D data into 2D computer graphics does not achieve the same graphical capability in relation to systems that use xyz coordinates (Jones *et al.*, 2008), which is needed to fully appreciate the true three-dimensional structures. Although the user has the capacity to gain more information from the dataset,

this is denied through the display choice. Two-dimensional displays also cause a reduction in dimensionality (Himmelsbach *et al.*, 2010) and depth, making it hard for the user to pick out the 3D forms that are within the point cloud (Kreylos *et al.*, 2008, p. 847). Additionally, remotely-sensed data, including laser-scanned data, are typically displayed and manipulated in a Geographical Information System (GIS), but these were initially designed for top-down 2D visualisation (Döllner and Hinrichs, 2000). In contrast to the typical user interface, it is argued that it would be more natural to interact with data derived from the 3D real world in three-dimensions (Faust, 1995). Ultimately, a 2D interface makes lidar analysis inadequate when considering the spatial relationships between structures, as Trinks *et al.* (2006) describe in their geological study of photorealistic virtual outcrops. This reported loss of spatial perception is denied from the viewer during the visualisation stage, which occurs in the highlighted area in Figure 2-5.

The visual inspection of point clouds taps into powerful human ability of pattern recognition (Kovač and Žalik, 2010; Brodu and Lague, 2012). In the case of Haala and Kada (2010), human operators were needed to identify intricate building façades complex forms from lidar point clouds during automatic 3D building reconstruction. Those who currently use 2D displays to view these 3D datasets can use pictorial cues, as was explained in the Introduction, and are able to manipulate the viewpoint to bypass the issue of occluded points. However, humans possess an in-built analytical capability that is currently underexploited.

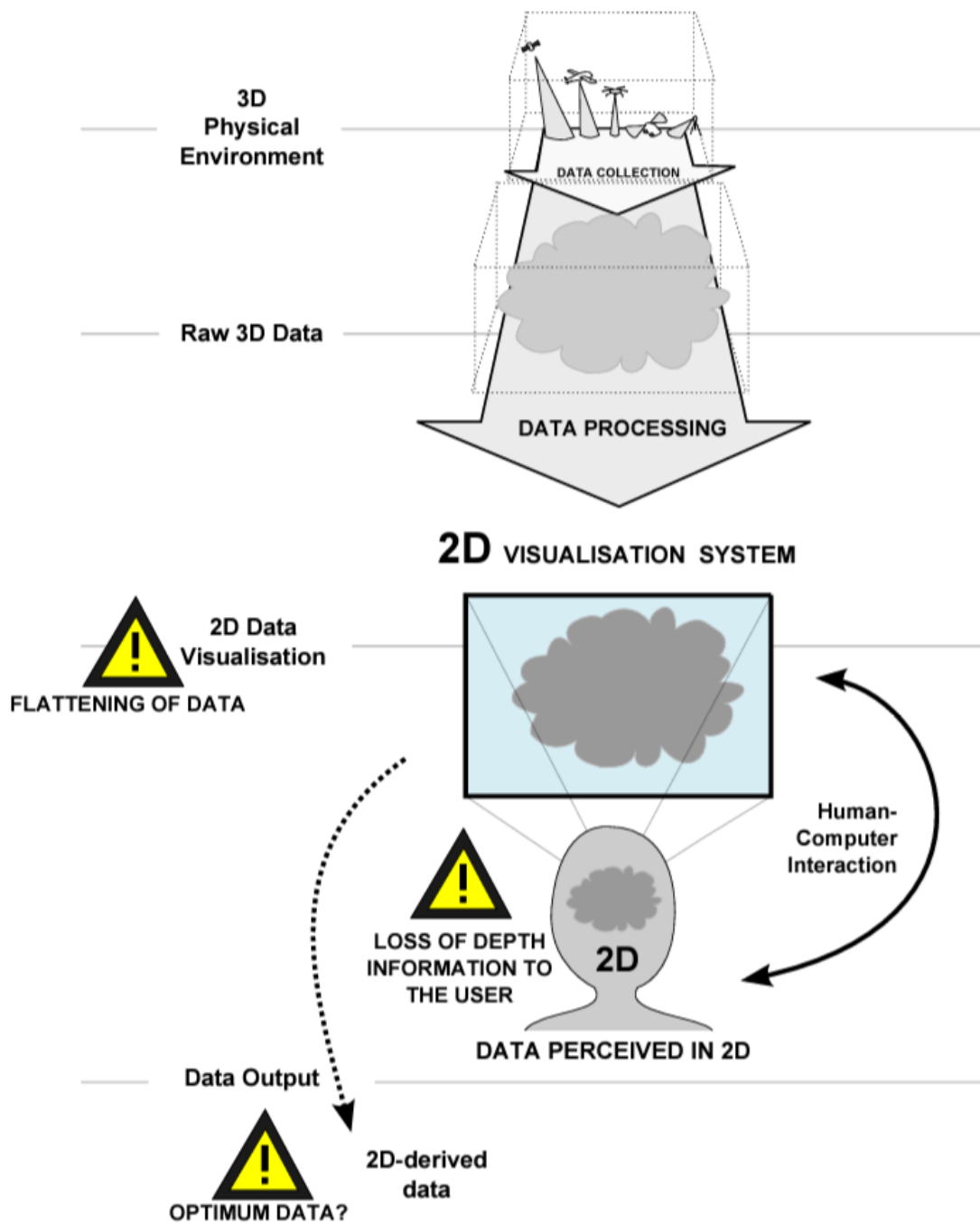


Figure 2-5. From lidar data acquisition of the 3D physical environment, to 2D visualisation of the data. The 3D data are flattened – does this affect visualisation outputs? Diagram author's own.

2.2.4 Proposed 3D visualisation

Discrete lidar points are inherently 3D (Warner *et al.*, 2003) because of their xyz structure – each lidar hit has a positional value (*xy* coordinate) and a *z* value for its height. These three dimensional values place it in 3D space. A 2.5D lidar visualisation study by Gardner *et al.* (2003) states that the “*full spatial complexity*” of the raw point cloud is at the disposal of the human analyst (Gardner *et al.*, 2003, p. 30). With this in mind, Trinks *et al.* (2005) feel that it would be ideal to then take this 3D information and display and analyse it in three dimensions.

The spatial complexity of lidar data lends well to a 3D projection (Warner *et al.*, 2003) and point cloud data exhibit differences in accuracy (Haala and Kada, 2010), coverage, and spatially biased density (Nebiker *et al.*, 2010). Working with raw point cloud data preserves data integrity and, according to Kreylos *et al.* (2008), enhances the accuracy of point selection for the extraction of features. As well as heterogeneous coverage of the scanned environment, data collection surveys and interference with the data returns typically affect the data coverage. The voids of information between points compound this complexity Nebiker *et al.* (2010). In this study, particularly in section 3.2, the number of points per metre squared (ppm²) is stated, which is a measurement of the point cloud density. The value represents the number of lidar hits in a given area of ground and gives an indication of the size of the gaps or voids between the points.

Observations can also be drawn from other fields, to help justify use of stereoscopic 3D lidar visualisation. Smallman *et al.* (2001, p. 51) claim that 2D leaves each dimension ambiguous. Jones *et al.* (2008) 2.5D “reduces graphics capability relating to fully 3D systems”. As an alternative to Figure 2-5, Figure 2-6 shows that the physical environment is three-dimensional and this structuring is carried right through to the human cognitive system when using stereoscopic displays. Fundamentally, the immersion is more in tune with the human senses (van Dam *et al.*, 2002). This creates a virtual environment, which gives the user a deeper understanding of the visualised environment than a standard 2D desktop display (Neves *et al.*, 1997).

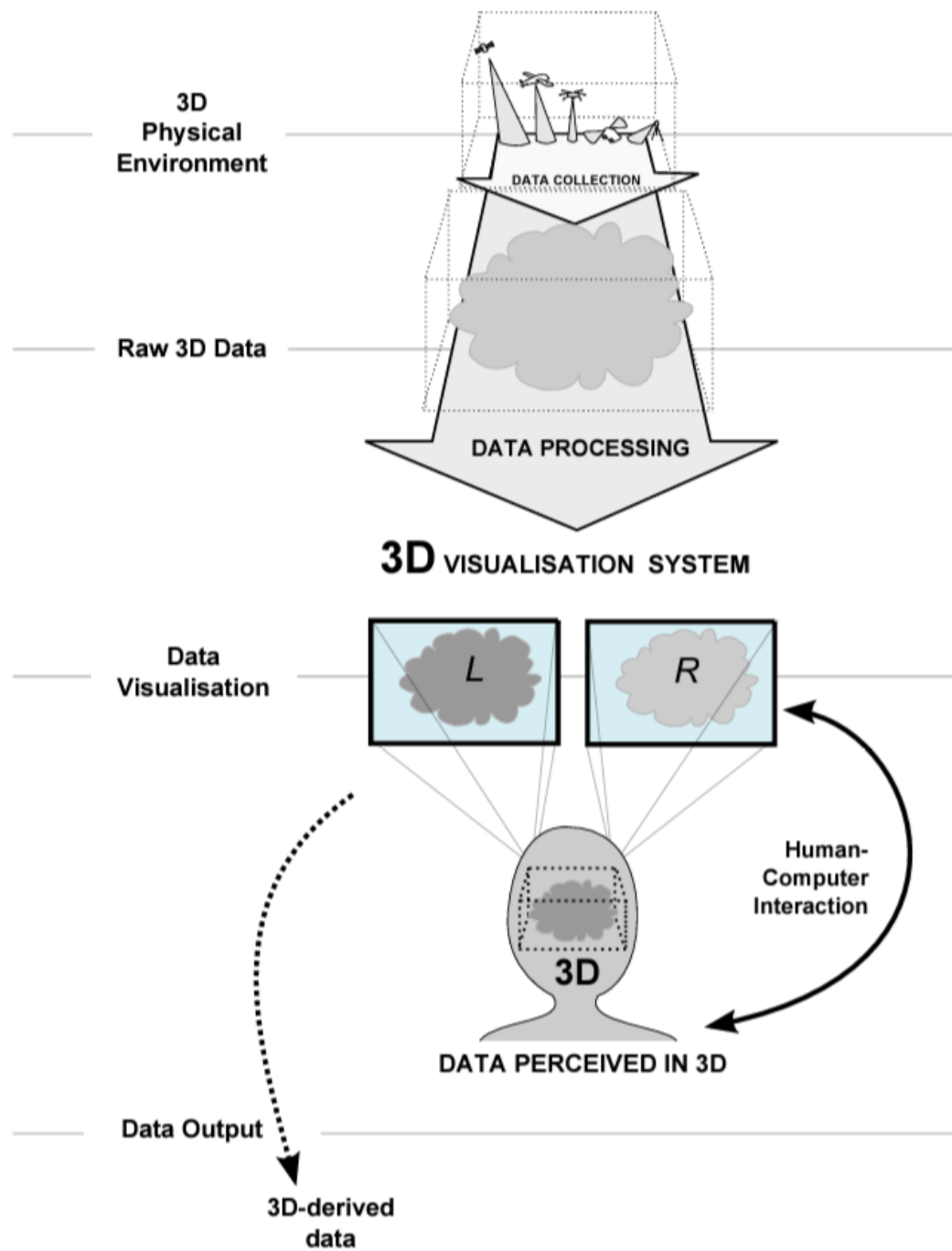


Figure 2-6. From lidar data acquisition of the 3D physical environment, to stereo 3D visualisation of the data. Could the data output be an improvement on 2D display-derived data? Diagram author's own.

Complementary 3D visualisation characteristics

Here, the advantageous characteristics of immersive data representation are highlighted to justify its exploitation in lidar analysis, in response to Kraak's (1998, p.192) question, "How does the use of an expedient such as a stereoscope influence the map reading task?". Firstly, the immersive nature of the visualisation system enables the user to experience (geo)virtual environment (Hodza, 2009), helping to create deeper knowledge of the viewed information (Neves *et al.*, 1997). Burwell *et al.* (2012), echo this sentiment, stating that stereo 3D arguably gives the users a sense of closeness to 3D structures, demonstrating this with respect to lidar point clouds. Figure 2-7, after Brodlie *et al.* (2002), shows how, from (a–d), users become increasingly immersed in representations of geographical environments as virtual reality increases. Figure 2-7(a) would describe a 2.5D interaction, where the user of a 2D screen is segregated from the data (standard lidar display), and from (b) to (d) the user becomes increasingly immersed in the represented information, feeling as if they were in the virtual world in (d). As the user becomes more immersed in the representation, theoretically, patterns in geographical data, such as point clouds, could be explored and analysed more effectively.

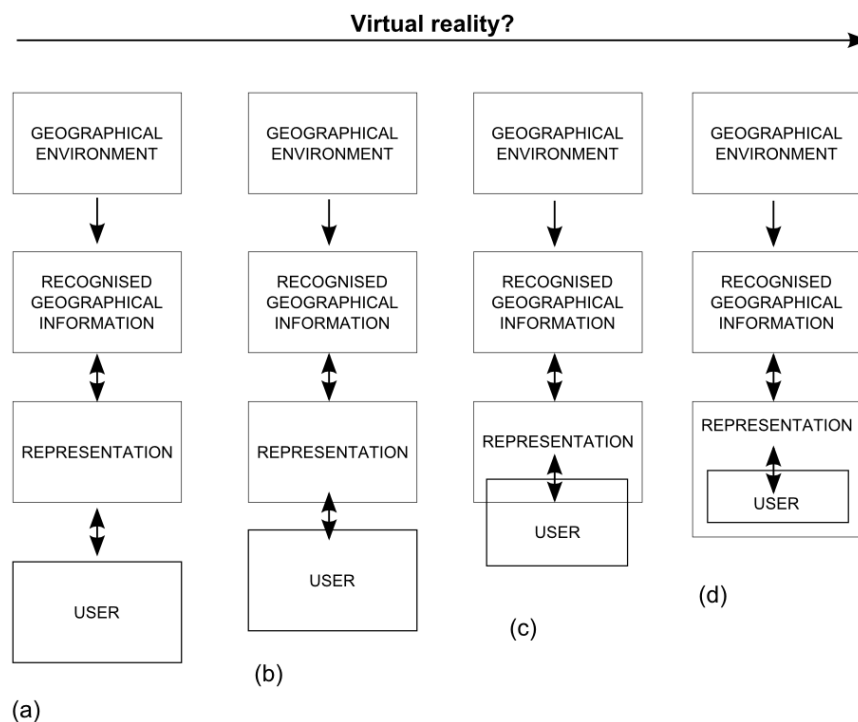


Figure 2-7. Different stages of immersion, after Brodlie *et al.* (2002)

The characteristics of 3D interaction include manipulation and analysis and McIntire *et al.* (2014) found that stereoscopic 3D displays were, alongside finding/identifying/classifying objects or imagery, most useful for tasks involving the manipulation of objects. When a user is within and interacting with the same environment as the data, this lessens the cognitive load on the user. This 'brain-strain' is higher when the user has to think, see, and control the data in contrasting dimensions (see Figure 2-8), e.g. using a 2D mouse on a flat surface to manipulate the 3D data represented on a 2D computer screen (Mark, 1992). With this in mind, the fewer cognitive constraints on the user, the better. This is especially true if the geovisualisation acts as a tool that assists "(visual) thinking process" (Kraak, 1998, p.12).

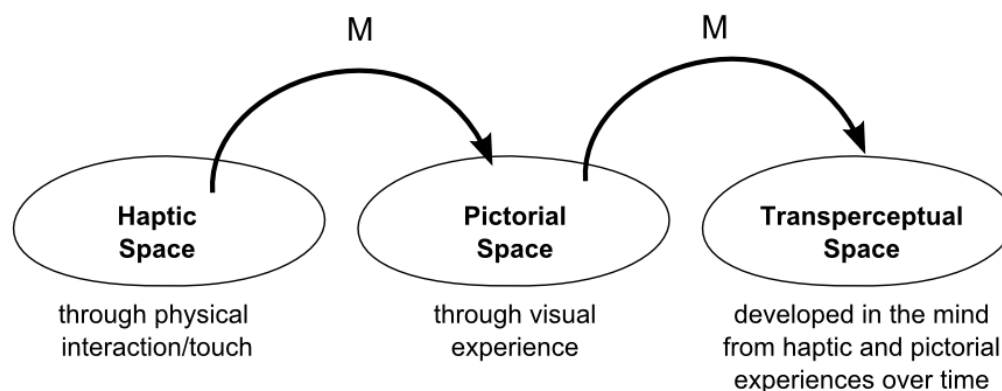


Figure 2-8. Three cognitive spaces, adapted from Mark (1992). M represents the 'metaphors or mappings' between the spaces. Transperceptual space is built from the experience of other spaces.

The freedom of movement that is permitted in an immersive display allows the user to change angles and avoid occlusions. Traditional cartographic methods use a top-down method whereby the Earth's surface is displayed from above (Faust, 1995), but, users of alternative virtual data exploration have found their new-found ability to fly over elevation and freedom to 'trespass' advantageous (Hodza, 2009, p. 516). Although ALS is acquired from a top-down perspective, the point cloud structure permits alternative viewing perspectives. In 2D, data points can be hidden behind one and other, which hinders communication of information from the map to the user (Kraak, 1993, p. 193). In a 3D environment, this occlusion is avoided if the user can navigate around the data, providing there are sufficient data. Three-dimensional models are considered to

allow researchers to observe different scales or scenarios that are inaccessible because of remoteness or danger (Lai *et al.*, 2010, p. 222). This virtual exploration is especially suited to remotely-sensed datasets, including those acquired from platforms such as UAVs that, as previously mentioned (e.g. Anderson and Gaston, 2013), increase the accessibility of study sites. This increase in access, with or without permission, also relates to areas of human and political geography research concerned with the exploration of and engagement with off-limits urban architecture (Garrett, 2014).

Previous 3D Point Cloud Visualisation Work

A handful of literature explores the 3D visualisation of lidar point clouds (Warner *et al.*, 2003; Kreylos *et al.*, 2006; Kreylos *et al.*, 2008; Bernardin *et al.*, 2011; Gertman *et al.*, 2012) . New knowledge in this field was added by the author and others (Burwell *et al.*, 2012) , who assessed the potential for 3D visualisation of lidar point cloud via a head-mounted display (HMD). The portability of the head-mounted display meant that it could be taken to lidar experts at a GIS mapping company and an international remote sensing conference (Remote Sensing and Photogrammetry Society Annual Conference 2009). The volunteers were presented with a point cloud of a coniferous tree, which was displayed in stereo 3D. One volunteer stated that they could not see a benefit to using the stereo 3D display over a 2D one, whereas another participant commented that 3D allows better understanding of the texture and detail of the presented vegetation point cloud (Burwell *et al.*, 2012) . Although valid remarks, the study did not compare the 3D display with standard 2D representation.

One early study by Warner *et al.* (2003), duplicated in Gardner *et al.* (2003), uses a 2.5D approach for lidar data analysis. Although the authors compare non-immersive raster-based and a vector-based software packages, Warner *et al.* (2003) recommend several virtual tools that can equally be applied to lidar point clouds within a virtual 3D environment. These include the creation, edition, and allocation of attributes to 3D vectors, and statistical analysis of spatial distribution. Kreylos *et al.* (2006) use a strongly visual approach to lidar

analysis, post-processing. The authors give geoscience researchers virtual tools to help answer visual-analytical problems through an immersive reality geovisualisation system. These can also be refined intuitively, by removing and adding points with a virtual brush. Metadata could also be added to the lidar data by attributing information, e.g. species, to the points or classifications. Kreylos *et al.* (2008) suggest that the standard visualisation and interaction methods used to segment point clouds is laborious.

Citing the Burwell *et al.* (2012) stereo HMD prototype, and others, Yan *et al.* (2015) predict that 3D city modelling would benefit from advances in data interpretation and exploration. However, this study anticipates that the application of 3D lidar visualisation could be more widespread, and effective, across the different areas of lidar analysis.

Gap 4: *There is a lack of lidar point cloud data represented in stereo 3D, despite its potential.*

Technological advancement

Recent virtual lidar display systems have included HMDs (e.g. Burwell *et al.*, 2012) and cave automatic virtual environments (CAVEs) (e.g. Kreylos *et al.*, 2006; 2008), the latter of which have projections on three to six sides. The resurgence in general interest in stereoscopic 3D technologies opens up the opportunity to test remote-sensing tasks in a variety of virtual environments. The technological advancement brings a reduction in nausea and a cognitive discomfort. Currently, lidar point cloud software is available as extensions of computer-aided design (CAD) packages, such as Terrasolid (Terrasolid, 2013b) and 3D animation software, and, increasingly, point cloud processing and visualisation software, according to Nebiker *et al.* (2010). In recent years, there has been a trickle of stereoscopic lidar analysis software onto the market that promote the ability for stereoscopic extraction of shapes directly from point cloud and basic measurement between points.

Open-source FUSION Lidar Data Viewer (LDV) software (McGaughey, 2014) has a 2D visualisation mode, which offers depth perception from horizontal

movement. The LDV applies continuous horizontal shaking movement of the data, named “wobble vision”, which adds motion parallax; as shown in Figure 1-3, the viewer gains some depth perception because of the fast movement of objects in the foreground in comparison to slower background objects. This was also noted by Garg *et al.* (2002), who studied 87 medical students’ knowledge of carpal bone anatomy following constrained vs. user-determined dynamic views of the 2.5D skeletal models.

The LDV also has an optional stereo anaglyphic mode, which allows the user to use red and green lenses to view the GUI with depth. However, VrLidar mapping software from Cardinal Systems (www.cardinalsystems1.net/, accessed 01-02-15), has an active polarised stereo view of data, but does not currently allow measurement of data interactively. Geovisionary software (www.virtalis.com/geovisionary/, accessed 03-05-15) uses active shutter glasses during user-navigation through of different data formats, including lidar, and permits basic measurement between points.

Although software packages exist that do apply human depth perception to the visualisation of lidar data, concept illustrated Figure 2-6 (a) there is not a widespread adoption of the method and (b) there is no evidence that compares the existing 2D display method against the immersive technique.

2.3 Evaluating geovisualisation methods

Proposed 3D techniques need to be compared against 2D because certain data might be sufficiently represented in 2D or 2.5D (Lai *et al.*, 2010; Neves *et al.*, 1997). For example, the visualisation of non-spatial data, such as a line graph, gains no added-value when projected in 3D (Lai *et al.*, 2010). With this in mind, a comparison of geovisualisation systems can be carried out to evaluate the strength and weaknesses of each (Fuhrmann *et al.*, 2005).

2.3.1 Participant evaluation

Discounting citizen science studies, which use members of the public for ground truthing and data analysis tasks (e.g. Foody *et al.*, 2015; Comber *et al.*, 2013), remote-sensing literature does not typically draw attention to human assessment of techniques. This may be because the recruitment of users raises some issues of bias, such as preference bias (Andre and Wickens, 1995) and evaluator effects (Hertzum and Jacobsen, 2001). Gardin *et al.* (2011) underline the absence of systematic methods to assess the effect of remote-sensing operator performance on data analysis.

Gap 5: *Despite the integral use of humans in the process chain, remote sensing visualisation techniques are rarely qualitatively evaluated*

Since comparison between prototypes and existing data analysis methods is not apparent in the lidar literature, relative examples are drawn from the wider geovisualisation domain. For the use of 3D visualisation technologies, the degree to which the user can see in stereo, also known as their stereoacuity, is an over-riding cognitive factor that pre-determines the application of the user's Geographic Expertise (GE). In this study, this is a factor that is taken into account in the methodology. Kastens and Ishikawa (2006) underline that it is a universal human ability to see patterns amid visual clutter. Furthermore, Robinson (2013) highlights that it is often assumed that HCI studies benefit from the analysis of expert users. The author queries when expert participation would be important and how much expertise is enough to make an impact on the outcome of an HCI study. This consideration is pertinent for visualisation studies based on performance differences in monoscopic and binocular depth cues, which are observed and not learnt. Evaluation can be achieved by understanding user performance or features of the visualisation system (Fuhrmann *et al.*, 2005; Kraak, 1998). But, crucially, Lai *et al.* (2010) ask, "*What are the criteria for measuring an effective representation?*" (Lai, 2010, p.233). Accuracy was frequently found to be the main, and sole, performance measurement of 184 2D vs. 3D visualisation experiments, reviewed by McIntire *et al.* (2014), which were from a range of applications (medicine, HCI,

engineering). Many of the studies also coupled accuracy with response time (McIntire *et al.*, 2014). Qualitative responses (participant preference) are considered to be important to the evaluation of visualisations, especially so, insist Andrew and Wickens (1995), when differences between user performance and preference are observed.

Gap 6: Which evaluation method is best for quantifying user performance?

2.3.2 Existing vs. prototype

Geovisualisation prototypes have been technically developed, but their assessment could be taken further. Fuhrmann *et al.* (2005) question whether novel tool designs are actually usable and useful for knowledge discovery and decision-making. However, past research has not made direct comparison between the proposed 3D and current 2.5D geovisualisation techniques (Neves *et al.*, 1997; Feng *et al.*, 2007; Burwell *et al.*, 2012). Other examples include a VR study by Mancera-Taboada *et al.* (2011). The research aimed to create a virtual environment for disabled users, who might otherwise be unable to access the featured architecture, but the prototype was not tested on users. Elsewhere, Feng *et al.* (2007) confidently state that their featured 3D visualisation technique gives accuracy assessment that is above and beyond the standard 2D co-localization technique, but the authors do not substantiate this claim by evaluating how it is more accurate. Zehner (2010) argues there is a need for 2D display windows within a VR screen to display text, maps, graphs regarding borehole locations and stratigraphy. However, the visualisation setup was not formally evaluated by users because they reportedly kept getting lost in the VE. Kreylos (2006) found that they, the authors, and the majority of other people that tried their immersive virtual reality lidar system were able to instantly perceive objects in three-dimensions and could accurately interpret collective points as features. Although this may be true, the study does not assess task performance in an unbiased manner; potential applications cannot rely on the say-so of the authors and their colleagues. Admittedly, visualisation assessments using qualitative measures could only be used as a benchmark

because the utility is relative to the requirements of the end user. Furthermore, it is difficult to measure utility of visualisation because the results are in the mind of the user (Gahegan, 1999). In any case, lack of prototype evaluation, either qualitative or quantitative in nature, means that the validity of the findings presented in the literature is uncertain. Humans are integral to the assessment of a geovisualisation technique and, therefore, provide the ideal tool for directly testing its effectiveness (Gahegan, 1999).

Gap 7: *Comparison between the prototype visualisation technique [3D] and existing [2D] is required when proposing a new method.*

2.4 Review outcome

The gaps identified in the literature review are shown in Table 2-2 and the schema in Figure 2-9 illustrates the problem statement that is drawn from the literature. 2D and 3D visualisation of lidar point clouds both appear to be viable methods, but the accuracy of their outputs is currently uncertain among the literary evidence.

Table 2-2. List of gaps from the Literature Review, numbered in order of appearance.

- GAP 1:** *DO NOT KNOW WHICH MANUAL LIDAR POINT CLOUD ANALYSIS TASK(S) PERFORM BETTER IN 3D, COMPARED TO 2D.*
- GAP 2:** *WHICH LIDAR VARIABLES ARE BETTER REPRESENTED IN 3D, AS OPPOSED TO 2D?*
- GAP 3:** *THE GEOMETRY OF THE DATA ACQUISITION AREAS MAY AFFECT VISUAL 2D AND 3D LIDAR VISUALISATION ANALYSIS OUTPUTS.*
- GAP 4:** *THERE IS A LACK OF LIDAR POINT CLOUD DATA REPRESENTED IN STEREO 3D, DESPITE ITS POTENTIAL.*
- GAP 5:** *DESPITE THE INTEGRAL USE OF HUMANS IN THE PROCESS CHAIN, REMOTE SENSING VISUALISATION TECHNIQUES ARE RARELY QUALITATIVELY EVALUATED*
- GAP 6:** *WHICH EVALUATION METHOD IS BEST FOR QUANTIFYING USER PERFORMANCE?*
- GAP 7:** *COMPARISON BETWEEN THE PROTOTYPE VISUALISATION TECHNIQUE [3D] AND EXISTING [2D] IS REQUIRED WHEN PROPOSING A NEW METHOD.*

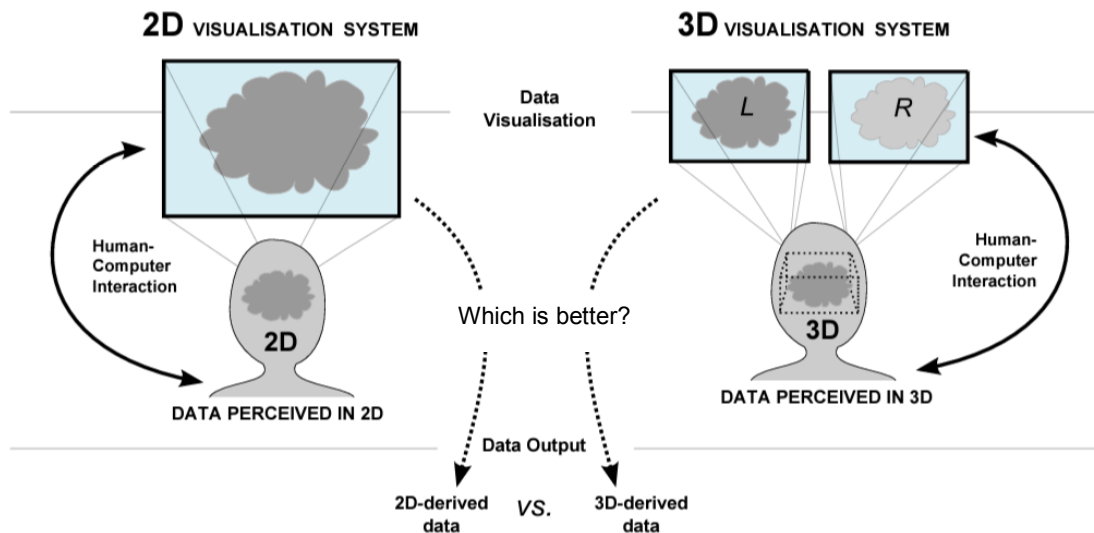


Figure 2-9. Visual summary of problem statement – which is more precise/accurate, with respect to lidar point clouds, data output from 2D or stereoscopic 3D visualisation? Here, 2D refers to 2.5D representation of (2D or 3D) objects on a flat projection, whereas offset projections of the same 2D/3D object to each eye result in viewer visualising the object in 3D space.

2.4.1 Research questions & hypotheses

The Research Questions, around which the study is developed, are listed below. They are based around the assessment of two manual tasks – measurement and interpretation, and assess the appropriateness of the novel method. A summary of hypotheses for each research question are listed below, which broadly assume stereo 3D visualisation adds value to manual lidar tasks by generating better accuracies than 2D. These are revisited in Chapter 8.

Table 2-3. Research questions and hypotheses

| Chapter(s) | Research Question | Hypothesis |
|-------------------------------------|--|---|
| Measurement CHAPTER 5 | <p>RQ 1 - Is there a significant difference in linear measurements of lidar point cloud features derived from 2D and 3D visualisations?</p> <p>1.1 – <u>How precise are point cloud measurements</u> made in 2D, in comparison to those made in 3D for (a) a planar feature and (b) a volumetric feature?</p> <p>1.2 - <u>Is there a significant difference between point cloud measurements</u> made in 2D, in comparison to those made in 3D for (a) a planar feature and (b) a volumetric feature?</p> | <p>H1 - There is a significant difference in linear measurement of lidar point cloud features derived from 2D and 3D visualisations for (a) a planar building feature (b) a volumetric vegetation feature.</p> <p>H1.1 - For (a) and (b), the range of measurement estimations will be smaller in 3D, compared to 2D.</p> <p>H1.2 - For (a) and (b), there is a significant difference between point cloud measurements made in 2D, in comparison to those made in 3D.</p> |
| Interpretation CHAPTER 6 | <p>RQ 2 - Is lidar point cloud feature interpretation more accurate when carried out via 2D (monocular) or 3D (biocular) visualisation?</p> <p>2.1 - For point cloud features of varying geometries, (a) what is the 2D and 3D interpretation accuracy? (b) is there a <u>significant difference between interpretations</u> of the point cloud features made in 2D, in comparison to those made in 3D?</p> <p>2.2 - Does the <u>complexity of the acquisition environment and subsequent processing</u> affect 2D vs. 3D accuracy results (a) on flat terrain and (b) on sloped terrain?</p> | <p>H2 - Lidar point cloud feature interpretation is more accurate when carried out via 2D (monocular) or 3D (biocular) visualisation.</p> <p>H2.1 - For point cloud features of varying geometries, (a) the interpretation carried out in 2D and 3D will vary according to the point cloud features. (b) The accuracy of interpretation of a point cloud feature is significantly better for 3D-derived classifications, compared to 2D.</p> <p>H2.2 - The complexity of the acquisition environment and subsequent processing affects 2D vs. 3D accuracy results. On flat terrain (a), interpretation is straightforward and on sloped terrain (b), interpretation is difficult, less so in 3D because of the added depth.</p> |
| Method Review CHAPTERS 5, 6, & 7 | <p>RQ3 - How effective is the methodology at comparing 2D vs. 3D visualisation of lidar point clouds?</p> <p>Consider the following: (3.1) measurement task, (3.2) interpretation task, and (3.3) general experiment.</p> | <p>H3 - The method used for the set-up and execution of the measurement task (H3.1), interpretation (H3.2) task, and the general experiment (H3.3) is suitable, but some improvements may be identified.</p> |

3.

Method

3. Method

3.1 Introduction

The structure of the method chapter echoes the chronological order of stages in the research project's timeline, outlined in Figure 3-1. The development stage of the research experiment occurred from February 2012 and lead up to the pilot study, main experiment, and subsequent data processing and analysis.

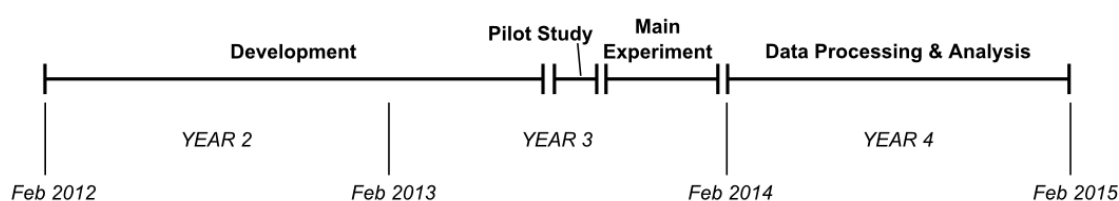


Figure 3-1. Timeline of the different stages of the study's methodology, following the initial year of literature review and project development.

3.2 Development

The initial period of development included (1) lidar data preparation, (2) visualisation system development, and (3) the design of the main experiment. Although this chapter presents these phases separately, in sections 3.2.1 to 3.2.3, in practice there were feedback loops between the three stages.

3.2.1 Lidar data preparation

Figure 3-2 displays the steps required during the data processing – airborne acquisition, classification of features, and the generation of the final data files for each area of interest.

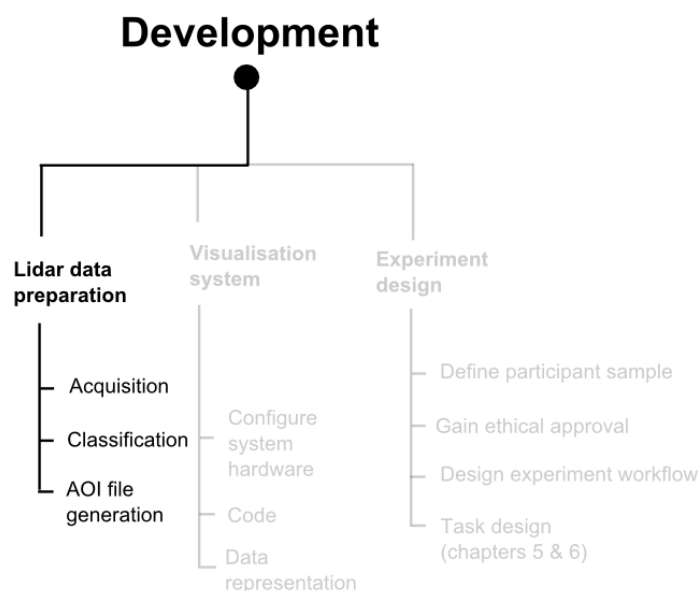


Figure 3-2. Node to show the 3 aspects of experiment development, with emphasis on the first lidar data preparation stage.

Acquisition

The airborne lidar dataset used in this research was issued cost-free by Airbus Defence and Space Ltd. (2013a and 2013b). The motivation behind the original ALS flight survey was to collect high resolution lidar for the generation of a commercial 3D urban modelling product. The data request was for a raw, unclassified product in areas where automated classification flowline had failed during previous commercial projects and had required intensive visual data quality checks by human data analysts. The locations of the acquisition sites, in Bristol and London, UK, are shown in Figure 3-3. The coverage concerns urban areas, owing to the city-modelling objective of the data collection company. The London data were acquired on 04-Jul-08 and the Bristol data were flown on 25-Nov-08, both using a plane-mounted Optech Gemini sensor (Optech, 2008).

Table 3-1 details the flight settings, as reported by Airbus Defence and Space Ltd. (*pers.comm.*, 2015). The London sensor configuration was set to a more acute acquisition angle (7° , rather than 10° used for Bristol), which would, in theory, create a higher point density.

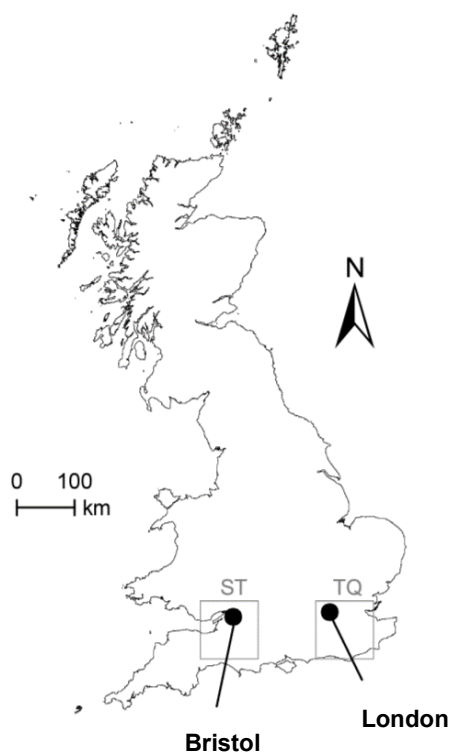


Figure 3-3. Map showing the locations of the laser-scanning acquisition sites in Bristol and London, UK, within Ordnance Survey (OS) national grid tiles ST and TQ. Ordnance Survey data © Crown Copyright and database right 2014.

Table 3-1. Specification of airborne laser-scanning (ALS) datasets acquired over Bristol and London study sites

| Data characteristics | Bristol, UK | London, UK |
|--|--------------------------------------|-------------------------|
| Tiles (OS grids) | ST565725 and ST575725 | TQ240800 |
| Size of tiles | 1km ² | 200m ² |
| Approximate locations | Clifton Bridge & Bristol City Centre | Shepherd's Bush, London |
| Acquisition Date | 25-Nov-2008 | 04-Jul-2008 |
| Sensor | Optech Gemini | Optech Gemini |
| Point Repetition Frequency (khz) | 125 | 125 |
| Scan Frequency (Hz) | 80 | 80 |
| Angle of acquisition (from nadir) | 10 | 7 |
| Points per metre-squared (ppm ²) | ~8 | ~10 |
| Raw format | .bin | .bin |

An initial aim of the study was to investigate the effect of different point cloud densities on 2D vs. 3D visualisation tasks, however there was little difference in density between the two datasets (between 8 and 10 points per metre squared or ppm²). The London and Bristol data were therefore treated as having a similar density and were processed using the same methods.

The use of this data justifies the visual approach used in this study and is a direct acknowledgement to literature that states certain environments that are difficult to classify with automated approaches (Meng *et al.*, 2010). Owing to the historical acquisition of the ALS surveys, the field sites were not visited for ground truth data verification – a significant time step of 5 years had passed between the acquisition and the visualisation project. However, in theory, a ground survey campaign timed with a flight survey, followed automated classification to highlight areas of ambiguous classification could be achieved. Practically, during a flight campaign, the features such as vegetation will not remain static for ground measurements and the inherent assumptions associated with ALS still remain (Zimble *et al.*, 2003).

This section of the method chapter focusses only on the practical processing steps that lead to the resulting .txt files that are used in the visualisation system.

Figure 3-4 shows a flowchart of the workflow, linked to each software package used to generate the final AOIs that are used in the experiment. The lidar data had undergone standard pre-processing steps prior to classification of the data by the researcher. The raw point clouds were issued to this project as .bin tiles of lidar coverage, the lidar format used in the commercial software Terrasolid (Terrasolid, 2013a; Terrasolid, 2013b) used by Airbus Defence and Space Ltd. The researcher was permitted to access to this commercial software during visits to the Airbus Defence and Space Ltd., which is where the automated classification process was undertaken. Once this stage was completed, subsequent processing was run on software and ArcMap, which was available via an academic licence. Table 3-2 lists each package used in the lidar processing workflow that is shown in Figure 3-4.

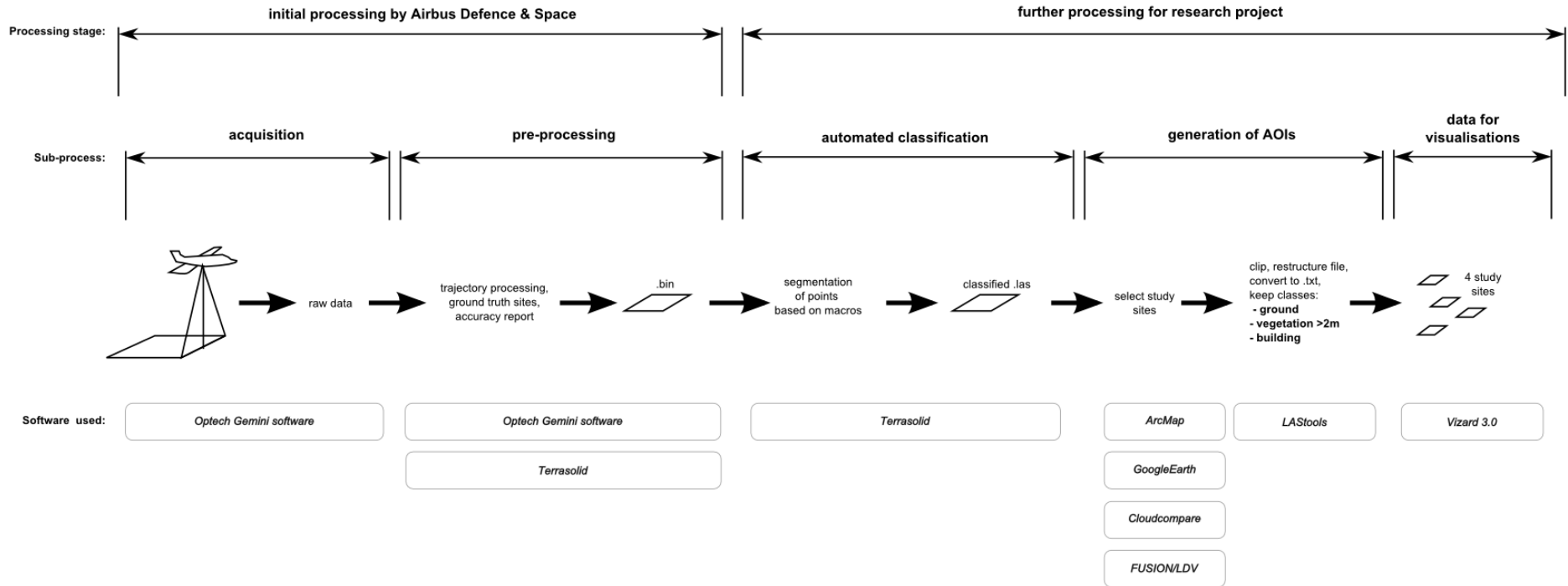


Figure 3-4. Flowline of lidar data processing and the software (detailed in Table 1-2) used to carry out each stage. Acquisition and pre-processing stages were carried out by Airbus Defence and Space Ltd., during 3D urban modelling product generation for UK cities, in 2008. The remaining processing was undertaken during this research project, in 2012.

Table 3-2. List of open-source and commercial software packages used during data editing for AOIs and visualisation.

| Processing stage | Software | Functions used | Source |
|---|---|---|--|
| Pre-processing | Optech Gemini | Trajectory processing, ground truthing against Ground Control Points (GCPs), accuracy reporting | via Airbus Defence and Space Ltd. |
| Pre-processing and automated classification | Terrasolid, MicroStation v8i (SELECTseries 2) | Classification of raw lidar, save as .las format | (Terrasolid, 2013a) (Bentley Systems Inc., 2010) |
| AOI generation | Cloudcompare | Verify extent of AOIs. Selection of POIs. | (Girardeau-Montaut, 2014) |
| | FUSION/LDV | Verify extent of AOIs | (McGaughey, 2014) |
| | ArcMap | View orthophotos, create AOI shapefiles | (ESRI®, 2014) |
| | GoogleEarth | Use Google Streetview as a ground truth guide during Interpretation methodology | (Google, 2014) |
| | LAStools | Convert .las tiles to .txt, clip AOIs to size according to min xy max xy, | (Isenburg, 2013) |
| Visualised data | Vizard 3.0 Vizard Lite 3.0 | Check data points would be acceptable for viewing power/frame rates. | (WorldViz, 2010b; WorldViz, 2010a) |

Automated classification

The original data were made up of several returns, or points, where the laser pulse signal has bounced off an object and wholly or partly returned to the sensor. Figure 3-5 shows the first, second, third, and last returns that have hit parts of a tree. The first return, 1, occurs when the laser first hits an object. Returns 2 and 3 may occur if the original beam is able to penetrate the feature. The 4th return is the last return. In the case of an impenetrable feature,

illustrated by the building (right), the first and last return occur at the same position. The data points used in the experiment are derived from the first returns only; second and third returns were ignored in the Bristol data. The London data exhibited noticeable intermittent gaps between some points (shown in Appendix A), so these data were reprocessed using second and third returns.

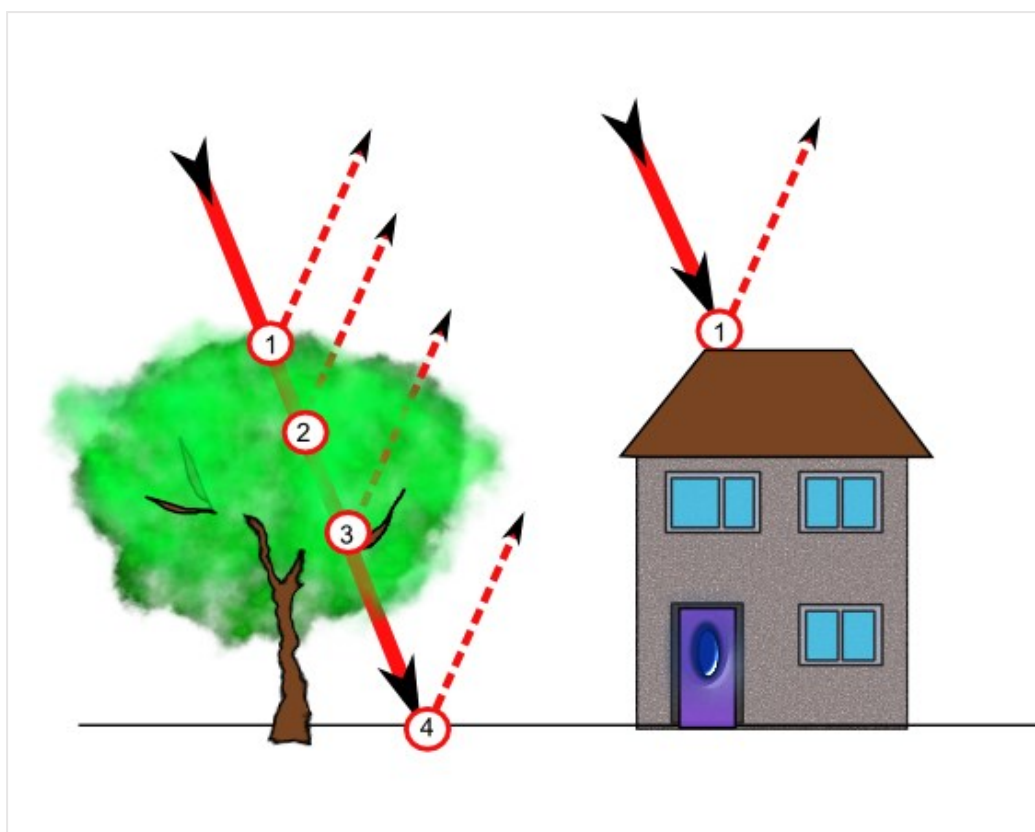


Figure 3-5. Schema showing the positions at which airborne lidar returns occur. The solid line denotes an in-coming laser that has been fired from an airborne instrument. Image author's own.

A classification routine was applied to the datasets using the TerraScan lidar-processing application of Terrasolid software (Terrasolid, 2013a), which was used in MicroStation v8i (SELECTseries 2) editing suite (Bentley Systems Inc., 2010). The macro, `classify_ht_bldg.mac` (code shown in Table 3-3 and Table 3-4), is made up of two steps, the first of which is shown in the first row of Table 3-3. This initial routine categorises the point cloud into either ground or non-ground by checking the points against the lowest point in a set area (refer to

pseudocode in Table 3-3). The points that meet the specified angle and distance from the original ground point are added to a triangulated irregular network (TIN) ground surface model. This process reduces the number of unnecessary points that are carried through to subsequent digital elevation model (DEM) generation that occurs after lidar processing. Consequently, the density of the classified ground points is much reduced in comparison to the raw data. The non-ground points that are positioned 2m above the DEM are held in a non-ground “vegetation” class, until the second step of the macro is run.

Table 3-3. Ground Points Classification Routine (Macro Step 1), which conditionally allocates the lidar points to ground and a (temporary) vegetation class. A pseudocode explanation is also provided. Numbers in brackets refer to the non-standard lidar classification of the ALS sensor format. *For Scene C, classification is also run with source classes set at second and third.

| | |
|----------------|---|
| Macro Step 1 | FnScanClassifyHgtGrd(5,100.0,2,6,2.000,999.000,0) |
| Input | Unclassified point cloud |
| Pseudocode | <ol style="list-style-type: none"> 1. Using the ground classification (5) as the point class into which ground points will be classified, 2. the maximum TIN triangle length is 100.0m, 3. take the first returns* (2) as the <u>source class</u>, 4. and allocate them to “Vegetation” (6) <u>target class</u> 5. if above a minimum <u>2m above the ground</u> 6. and below a maximum 999m above the ground. 7. Run this process on <u>all points</u> (0 fence). |
| Output Classes | Ground (5) + “Vegetation” (6) |

See Soininen (2005) for further information on Terrascan classification routines.

The ground and non-ground “vegetation” points undergo a second routine, which is used to extract a building classification from the non-ground points. The second step of the classification routine is detailed in the top row of Table 3-4, which contains pseudocode of the macro process and the set parameters. Unlike a routine that may thin the entire point cloud (regardless of the classification), the macro used in Terrascan (Terrasolid, 2013a) thins points on

the ground, but not other classes. The values for minimum building plane size and deviation from the roof plane were kept to the default commercial setting so that the point clouds in the experiment are representative of those presented to human operators in a production workflow. The Terrascan macro routines could be reproduced in other proprietary or open source software, using the parameters outlined in Table 3-3 and Table 3-4.

Table 3-4. Building Classification Routine (Macro Step 2), which further sorts the vegetation lidar points from the Macro 1 output into building and vegetation classes. A pseudocode explanation is also provided. Numbers in brackets refer to the non-standard lidar classification of the ALS sensor format. Z accuracy is deviation from the roof plane.

| | |
|----------------|--|
| Macro Step 2 | FnScanClassifyBuilding(5,6,15,3,40.0,0.20,0,0) |
| Input Classes | Ground (5) + “Vegetation” (6) |
| Pseudocode | <ol style="list-style-type: none"> 1. Using the ground classification (5) as <u>the point class into which ground points</u> have been classified, 2. take the high vegetation (6) as the source class from which to search building points, 3. and allocate them to building (15) <u>target class</u> 4. using normal rules (3) 5. if they have a minimum building plane size of 40.0msq 6. and have an elevation or Z accuracy of 0.20m. 7. Run this process on all points (0 fence). |
| Output Classes | Ground (5) + Vegetation (6) + Building (15) |

See Soinen (2005) for further information on Terrascan classification routines.

The macro processes in Table 3-3 and Table 3-4 generated data classifications that included building (class 15), high vegetation (class 6), and ground (class 5). These three classes are displayed to the participant during the experiment. Any points that were classified as the following were disregarded: unclassified, first, second, third, and last returns, low point, air points, low veg, bridge, overlap, and others. For details, refer to the Terrascan user manual (Soinen, 2005).

Generation of Areas of Interest (AOIs)

This section details only the processing steps taken to generate the AOIs (also referred to throughout as scenes) used during the experiment; the justification for the site selection is detailed in Chapter 5 and 6 (measurement and interpretation chapters). Four subset areas were clipped from the classified .las data files using `lasclip` and `las2las` from the LAStools suite (Isenburg, 2013). While creating subset areas of interest (AOIs), the .las data were converted to a comma-delimited text file structure. This created a format that was compatible with the visualisation software, Vizard 3.0 (WorldViz, 2010b), which is not a GI system and cannot read standard lidar file formats. The clipped AOIs were converted from .las into .txt format, using the `las2txt` LAStool function (Isenburg, 2013). An example of the file structure is `x,y,z,i,r,g,b,c`, where `xyz` = point coordinates, `i` = intensity of the return, `rgb` = red, green, and blue true-colour values derived from orthophotography, `c` = classification. Each line of the file shows these values for one individual lidar point, although only the positional `xyz` coordinates (first 3 fields) and the classification information (last field) are considered in this experiment. Fields 5-7 refer to RGB (red, green, blue) values that were matched to the lidar points using the TerraMatch extension of Terrasolid (Terrasolid, 2013b). However, these values, and intensity (4th field), were not used during visualisations. The reasons for their omission are discussed later, in section 3.2.2, during the development of the visualisation system design. The processing steps and LAStool command-line code used to create each of the four AOIs are shown in Appendix A. The four processed AOI text files provide the point cloud information that is carried through to the visualisation stage.

Data for visualisations

The AOI data files (.txt) are available to view in the digital appendix and the density of the resulting scenes is shown in Table 3-5. A comparison of the two ground columns in Table 3-5 highlights the effect of the macro classification routine (defined in Macro 1, Table 3-3) on all scenes. Appendix A shows density screenshots, which show raw *versus* processed lidar coverages.

There is a reduction of ground points from an original density of 5-10 ppm² to a processed density of <2 ppm², whereas the other processed classifications (vegetation and building) retained the majority of their original points. This is owing to the thinning process that the macro undertakes on ground points. In each scene, points are omitted from the visualisation if they are not classified as either ground, vegetation over 2m, or building. Evidence of omitted points can be seen in the Scene C dataset used in the experiment, Appendix C (refer to Scene C – point cloud data set, Plan View), which has clear gaps where cars are parked on the road.

Table 3-5. A measure point density tool (Terrasolid, 2013a) was used to analyse the density of the classified point clouds. Grey cells indicate absence of feature from a scene. Scene C vegetation was not measured.

| | BEFORE PROCESSING | | | AFTER PROCESSING | | |
|---------|---|------------|----------|---|------------|----------|
| | Density of raw point cloud (points per metre diameter) | | | Density of processed point cloud (points per metre diameter) | | |
| | ground | vegetation | building | ground | vegetation | building |
| Scene A | 8.900 | | 7.589 | 1.872 | | 7.130 |
| Scene B | 4.877 | 4.749 | | 1.464 | 4.609 | |
| Scene C | 10.301 | - | 11.154 | 1.999 | - | 10.886 |
| Scene D | 4.940 | 3.871 | 4.227 | 1.604 | 3.985 | 3.934 |

Table 3-5 highlights that there are differences in data coverage between the scenes that are visualised – the point cloud density is not the same across the four AOIs. However, densities of the point cloud densities for each of the scenes (A to D) are the same between the 2D and 3D visualisations, which allows for like-with-like comparison. Table 3-6 provides an overview of the datasets used during the visualisation experiment.

Table 3-6. Data files used during generation of Scene A, B, C and D datasets, using Terrascan software (Terrasolid, 2013a). The .txt files used in the visualisation system included ground, building, and high vegetation (over 2m). Lidar data © Airbus Defence and Space Ltd. (2013a & 2013b).

| Processing stage | | Scene A | Scene B | Scene C | Scene D |
|------------------|---------------------------------------|-------------------------------|---|--|--|
| Acquisition | ALS flight location | Bristol | Bristol | London | Bristol |
| | Acquisition angle | 10 | 10 | 7 | 10 |
| | Date | 25-Nov-2008 | 25-Nov-2008 | 04-Jul-2008 | 25-Nov-2008 |
| | Flightline overlap | Overlap | No overlap | No overlap | No overlap |
| Input | Raw lidar file | ST575725.bin | ST575725.bin | TQ240800.bin | ST565725.bin |
| Classification | Classified .las file | ST575725_edited_RGB | ST575725_edited_RGB | TQ240800_c123456_1 5normal | ST575725_edited_RGB |
| | Returns included in .mac routine | 1 | 1 | 1, 2, 3 | 1 |
| Output | .txt file used in visualisation | BEd26SceneA | BEd25SceneB | LHT1thc5615drop1234 | BCLF1thrgbcNO1 |
| | Min x,y (OSGB) | 357550.19, 172725.8 | 357840.01, 172582.0 | 524108.24,180159.9 | 356522.26,172913.24 |
| | Dimensions (m) | 29.78 x 33.7 x 15.1m | 43.96 x 37.92 x 25.26m | 52.89 x 34.49 x18.42m | 100.22 x 60.09 x 75.33m |
| | Dimensions stated to participants (m) | (~ 30m ²) | (~40m ²) | (~50m x 35m x 18m) | (~100m x 60m x 75m) |
| | No. of points in file | 2,305 | 4,396 | 10,628 | 14,574 |
| | Description | Single planar feature – house | Single volumetric feature – canopy of trees | Flat, suburban area, with planar and volumetric features | Sloped, valleyside area, with planar and volumetric features |

3.2.2 Visualisation system development

Figure 3-6 summarises the steps taken during development. The visualisation system was built in a Vizard Virtual Reality Toolkit 3.0, using the developer and Lite versions (WorldViz, 2010b; WorldViz, 2010a). The software used is commonly coded to carry out psychology experiments in a virtual environment. However, here, the software is employed to allow users to view the lidar scenes stereoscopically (or in 2D), and to tailor the human-computer interaction to the specifics of the research questions. Coupled with the necessary hardware, the VR software allows stereoscopic rendering of objects within a bespoke virtual environment.

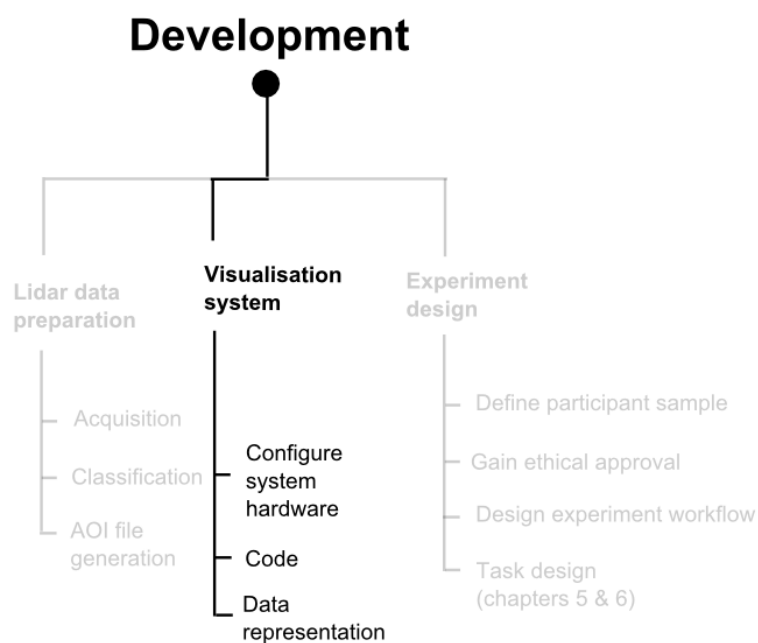


Figure 3-6. Diagram highlighting the considerations of visualisation system development.

System hardware

The apparatus used for the visualisation is listed in Table 3-7, with images of the 3D shutter glasses and gamepad controller shown in Figure 3-7. Xpand X102 Active Shutter 3D (www.xpand.me, accessed 01-04-15) were worn by participants when viewing 3D visualisations. The battery-powered glasses flicker each lens on and off, a rate that is too fast to be noticed with the naked eye. While the left lense is closed, the right lense is open, and *vice versa*.

The direct handling of geographical objects is influenced by the available 3D interaction technique (Döllner & Hinrichs, 2000) and Kreylos *et al.* (2006) state that the selection of heterogeneously structured sets of 3D points poses a challenge. The use of a computer mouse can hamper creativity (Mitasova *et al.*, 2012) because of Mark's (1992) cognitive metaphor issues. In light of the literature, a Saitek gamepad (Saitek, 2007) was chosen as the interaction device, Figure 3-7. Furthermore, during its use, there was relatively little lag between the gamepad and the PC response. When using a Kinect (Microsoft Corporation, 2015) device to detect user movement and comments, as Izadi *et al.* did during their 2011 real-time scanning for augmented reality (Izadi *et al.*, 2011), there was noticeable lag between user and on-screen movement. Also, the default skeleton structure used to track user movement was not sufficiently articulated, so only large, sweeping movements could be tracked. A Nintendo Wii remote (www.nintendo.co.uk, accessed 01-05-15) was also considered, but this too experienced a delay between use in the real-world and on-screen. The buttons on the Saitek device (Saitek, 2007) could be programmed according to the interaction requirements of each task. Its general purpose during user navigation was to provide movement of the body with the left analogue stick and movement of the head with the right analogue stick, as is later described in section 3.4.2.



Figure 3-7. Hardware used by participants in the experiment. Saitek gamepad (Saitek 2007), left, and Xpand X102 Acrive Shutter 3D glasses (www.xpand.me, accessed 01-04-15). Images author's own.

Table 3-7. Hardware required for visualisation set-up

| General description | Function | Model Used |
|----------------------------|--|---|
| Projector | Display the data onto a large screen | Christie Mirage WU7K (http://www.christiedigital.com/ , accessed 01-04-15) |
| Infra-red (IR) emitter | Synchronises the shutter glasses with the intermittent left and right images shown to the screen | |
| 3D shutter glasses | Shows the left eye one image and the right eye another image | Xpand X102 Active Shutter 3D (www.xpand.me , accessed 01-04-15) |
| Gamepad controller | Permits the user interaction with the displayed data - allows navigation and selection of points | Saitek P2900 (Saitek, 2007) |
| CPU | Computer processing unit used to run the visualisation | Windows 7 Enterprise, Intel® Core™ i7 CPU x 980 @3.33GHz, 64-bit operating system, 16.0 GB |
| Wireless mouse & keyboard | Allows the researcher to run commands / input information to the visualisation system | generic |

Code

Python-based (www.python.org, accessed 22-05-15) programming scripts were developed with guidance from a Vizard-specific coding manual (McCall, 2008). The coding script is broken down into separate modules, which are listed in Appendix B. The main modules, relating to the 2D and 3D visualisations of each Scene, call in the other modules and are listed in Appendix B. Copies of the scripts are included in the Digital Appendix.

Data representation

In a previous 3D lidar visualisation study (Burwell *et al.*, 2012), some novice users demanded interpolated data for contextualisation. However, Warner *et al.* (2003) claim that their 2D point cloud visualisation represented individual trees more clearly, in contrast to interpolated data. Additionally, users from the Burwell *et al.* (2012) study who were actively involved in vegetation and heritage lidar applications were interested in the raw 'true' representation of the

data and the potential to derive measurements from the points. For vegetation analysis, the assumption is that an increase in lidar hits implies there is an increase in leaf area (Sanz-Cortiella, 2011). Various aspects of the general capabilities of the visualisation are considered during its development.

Rendering. A minimum of 10 frames per second (FPS) and a mean of 24 FPS is required to be truly interactive, according to Gahegan (1999); the 2D and 3D visualisations were rendered at 60 FPS when no points were in view. When all points were in view for 3D the largest point cloud (Scene D, with 14,574 points), the slowest rendering speed was 0.60 FPS for 3D and 1.13 FPS for 2D. As with other VR geovisualisations, such as Hoang *et al.* (2010), the speed of the system inevitably declined with the rendering of more objects, although with constantly changing views, the rendering of the 2D and 3D visualisations for this study were acceptable. The computer performance dictates the size of the AOI and a higher number of points would mean a slower response between the user's actions in the real world, compared to the action on-screen. With regard to the CPU used in this experiment, refer to Table 3-7. The data representation is configured in the code, where shading, texture, perspective, and colour are applied to the data to, according to Kraak (1993, p.194), enhance 3D perception for both 2D and 3D visualisations.

Shape. The shape of each point cloud point needed to be three-dimensional. Although other lidar software, including GeoVisionary (Napier, 2011) and CloudCompare (Girardeau-Montaut, 2014) use a cuboid shape, by default, the lidar points here were displayed as spheres. Instead of being influenced by the harsh specular contrasts of a planar-faceted shape, the smooth, rounded point allows the eye to seek subtle shading differences according to a point's position.

Lighting and Shading. Two studies, Bernardin *et al.* (2011) and Sgambati *et al.* (2011), illustrate that dynamic lighting in a visualisation system can help to avoid losing topographic features in shadowed areas and can enhance pictorial depth cues. In an effort to keep 2D and 3D visualisations the same, participants do not have the option of changing the lighting of the scenes. Nevertheless, the visualisation examples underline the importance of lighting effects. In this system, the shading of the point cloud was caused by light diffusion from a top-down direction and from the viewer's position. This gave the points shading as if the user were shining a torch onto the points, which were already lit from a sun source.

Size. Point-size was determined by eye within the virtual environment so that (a) a sufficient gap was made between points to prevent converging, and (b) points were large enough to be seen at a distance. For Scenes A, B, and C, which were all acquired over reasonably flat terrain, scale of the points was set as `.scale(1.5,1.5,1.5)`, which translates to a 20cm diameter. For the sloped scene, D, the points were generated at a larger scale because of its larger z range, which stretches vertical distribution of points.

Colour. Participants in the Burwell *et al.* (2012) 3D VR trial stated that colouring of the lidar point cloud, which represented a coniferous tree, should be thematic (by height of classification) or realistic. Although the comparison of true colour point clouds in 2D and 3D was of interest, the aerial imagery effect of lean was evident in the imagery.

Figure 3-8 highlights the positional discrepancy between the aerial photograph and the overlaid point cloud for Scene B. It was assumed that the use of mismatched colouring with the lidar points would introduce confusion, which would be disorientating, particularly to those viewing a point cloud for the first time. Furthermore, the realism of geographical information may not necessarily enhance the data (Kraak, 2002) and can be a distraction from the task in hand (MacEachran, 1999).

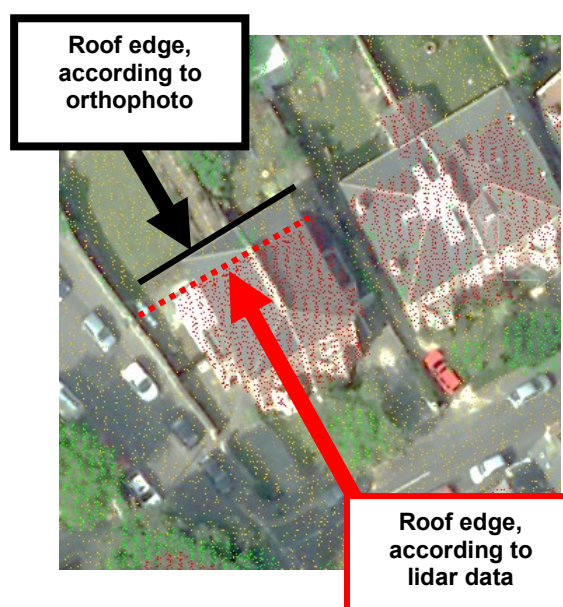


Figure 3-8. Screenshot from TerraScan software (Terrasolid, 2013), no scale, showing Bristol lidar data points (acquired 25-11-2008) overlying orthophotography (14-04-2007). Arrows highlight the impact of aerial imagery lean on true-colouring of lidar point cloud. Red points = building, green = vegetation, orange = ground. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophotograph © Geoperspectives (2013).

GUI Layout

As expected with any mapping interface, Gahegan (1999) insists that a North arrow is needed to help user orientation. On a stereoscopic interface, however, the presence of a 3D object close-up on the focal plane, *i.e.* the screen, is likely to cause discomfort when a user is focusing on a dataset ahead of them (Hoffman *et al.*, 2008). It was decided that the xyz axes should be highlighted by placing a wireframe bounding box at the edge of the lidar data volumes. Features were removed from the screen for the 3D test conditions, including a 2D plan view that allowed top-down orientation of the scene and 2D drop-down menus that offered different classification categories to users. Although this information could be useful for user orientation and sense of place, the “visual clutter” was removed (Sgambati *et al.*, 2011, p.40) to preserve the truly 3D experience and reduce any eye strain.

System Differences between 2D & 3D Methods

The 2D visualisation used the same software as the 3D display, but it was not projected under the stereo mode. Therefore, only the projector and screen were required for the hardware 2D display; the 3D glasses were not worn. The Vizard 3.0 software (WorldViz, 2010b) has some in-built commands for 3D settings, whose variables can be adjusted to suit the experiment set-up and user. These variables include screen distance, vertical field of view, and fusion distance value, which were kept constant for each user. The user's eye separation, or interpupillary distance, `,<viz>.ipd`, was tailored to each individual viewer to maximise the sense of stereoscopic depth information. According to Vizard 3.0 specifications, "*Values greater or less than the observers' actually eye separation will result in a decreased or increased sense of stereoscopic depth information, respectively.*" (WorldViz, 2010b)

During testing, it was revealed that the shutter glasses did not always correctly display the left eye's image to the left eye and the right image to the right eye. This only became apparent once the code had been launched, so although a negative IPD value could reverse the effect that too could only be set prior to running the main code under the stereo setting. This technical issue could not be fixed in time for the trials, so it instead, when required, the participants were asked to, unconventionally, wear the glasses upside-down. In theory, the scenes could be re-launched until the shutter glasses were correctly synchronized, but in practice this would leave participants turned away from the screen for a period of time. During the visualisations, the 3D depth was verified by the interviewer, who has a high stereoacuity of 20 secs of arc.

3.2.3 Experiment design

The experiment design process is highlighted in Figure 3-9 and considerations regarding the target participant population and ethical requirement are detailed here. However, the majority of the task design is detailed in Chapter 5 and 6.

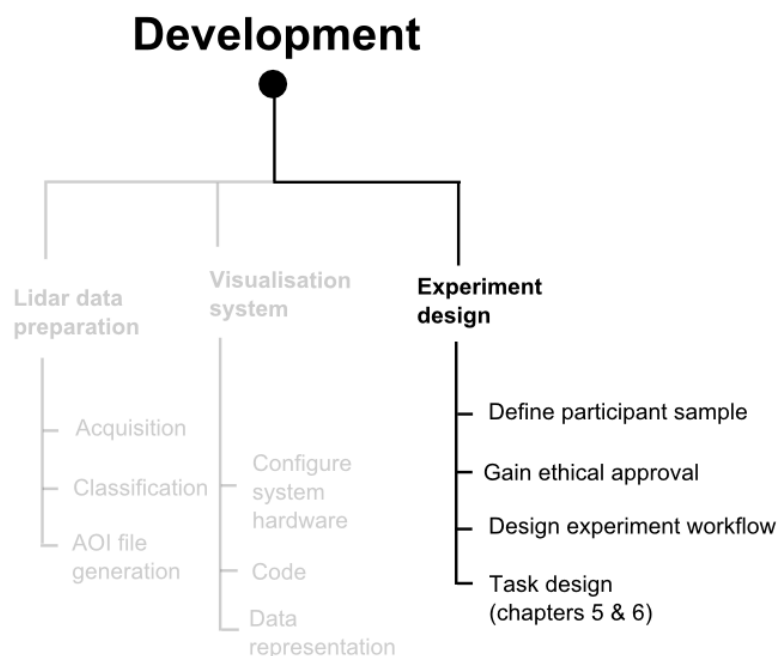


Figure 3-9. Diagram highlighting the experiment design aspect of the experiment development.

Defining the Participant Population

When developing a tool, there needs to be a specific user with a specific application for whom the tools are created. Similarly to the Burwell *et al.* (2012) study, an HMD could be used for 2D vs. 3D studies, permitting a portable experimental set-up. However, the headsets that were available during experiment development, the Emagin z800 (eMagin, Bellevue, WA, USA), have a restrictive 40° field of view (fov), which meant that the user cannot feel ‘immersed’ in the visualisation because the fov is not filled with the data. Instead, the availability of the Virtual Reality Theatre at the University of Leicester’s Geography Department, coupled with the high availability of university staff and students around the campus-based experiment had potential to generate a large sample size. As the study investigates the human ability of stereo vision and data, it is argued that stereo depth cues are exploitable in all user groups, whether they are expert or informed. A random dot stereo Randot™ stereo test (Stereo Optical Co., 2009) is used to measure the participants’ level of stereoacuity. The test, shown in Figure 3-10, is graded from 1 to 10 (400 to 20 seconds of arc at 16 inches), which is low to high stereoacuity.



Figure 3-10. Randot™ ('Random Dot') Stereo Test (Stereo Optical Co., 2009). The 'Circles (with random dot ground)' test, was carried out by each participant to determine their stereoacuity. 10/10 correct answers = 400 seconds of arc. As shown in the image, the other tests were obscured by cardboard. Image author's own.

It is commonly assumed that HCI studies require feedback from experts, and studies such as Edler *et al.* (2014) presume that participants with a geography background would have '*advanced expertise*' in map-reading (Edler *et al.*, 2014, p.157). The authors imply that this is advantageous over using non-geographers to take part in their stereoscopic map-reading study. However, an exclusive approach to participant selection may have consequences on the results. For example, Bleisch and Dykes (2014) recruited participants who were informed in geomatics and geography, but the authors state that those recruited may have been less informed than anticipated. Robinson's (2013) thought-provoking questions in relation to participant geographic expertise are relevant here; what is the best practice for categorising experts and at what point does one become an expert?

This study focuses on the interpretation of geographical features, but the experiment is designed so that participants do not need to be categorised according to expertise. The potential end users of geovisualisation techniques are anyone who has a reason to analyse lidar point clouds visually and

manually. The analysis is based on the general cognitive response to stereo and non-stereo stimuli in relation to geographical scenes, though the findings may be relevant to other applications outside this domain.

Ethics

The ethics assessment that was carried out for the trials, which outlined the implications of the study on the participant. Permission from each volunteer was sought through a Consent Form, which volunteers received prior to the experiment, along with the Information Sheet. These documents are available in Appendix G and were given to the volunteers in the main experiment and the pilot study, which was performed beforehand.

Experiment Workflow

The workflow of the main experiment was influenced by the target population, the visualisation system, and the available lidar datasets. The pilot study was a major aspect of the development, particularly because there is little evidence of literature that covers the design of a remote-sensing visualisation experimental design (Gardin *et al.*, 2011).

3.3 Pilot Study

A pilot study was carried out in October 2013 using a small number of volunteers, in order to explore the suitability of the presented lidar datasets, the capabilities of the visualisation system, and the experiment tasks. As shown in Figure 3-11, the findings lead to amendments to the visualisation system and streamlining of the experiment. Owing to the corruption of a hard drive that resulted in loss of data, some details of the pilot study are not fully reported. However, the pilot participant feedback is described here as fully as possible.

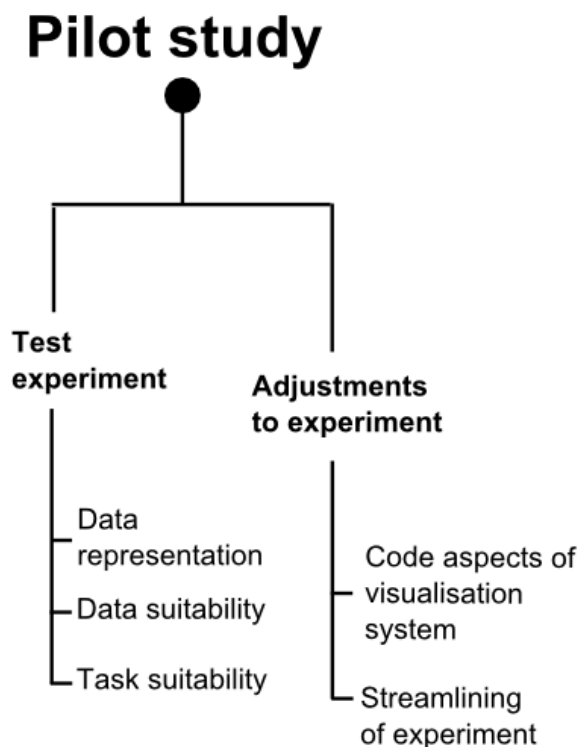


Figure 3-11. Elements of the pilot stage of the methodology.

3.3.1 Set-up of pilot

The primary aim of the pilot study was to test the suitability of the point cloud data and tasks that the participants would undertake during the main experiment. Most aspects of the pilot experiment were the same as the main experiment (section 3.4). The pilot was carried out in the same VR theatre as the main experiment, using the same set-up (section 3.4.2) and datasets (Appendix C). The same participant documents (Appendix G) and written survey (Appendix I) were used. Similarly to the main experiment flow, section 3.4.3 and Appendix F, two tasks were presented to the pilot volunteers.

The approaches were altered each time, based on the recommendations of the previous pilot participant. The iterative improvement and evolution of the experiment during the pilot process meant that a statistical comparison between the observations cannot be made.

3.3.2 Pilot participants

Four volunteers carried out tasks and provided invaluable feedback regarding several aspects of the experiment. It was essential to pilot the aspects of the experiment with volunteers whose lidar expertise varied because participant recruitment for the main experiment was not based on any prerequisite lidar knowledge or expertise. Therefore, the individuals were recruited because of the varying level of lidar knowledge and experience, which was already known to the researcher.

The participants were made up of 3 males and 1 female, all aged within the range of 26-35. The stereoacuity of each of the volunteers met the required standard (*i.e.* >5/10 on the Randot™ Stereo Test, Stereo Optical Co., 2009), as summarised in Table 3-8 (PILOT1 and PILOT3 stereo results not available). The level of expertise was derived based on self-reporting, on a scale (none > capable > proficient > highly proficient > expert).

Table 3-8. Summary of basic background characteristics of the four pilot participants. Some characteristics, e.g. nationality, are omitted from the table either because of data loss or to preserve the anonymity of the volunteers.

| Participant | PILOT1 | PILOT2 | PILOT3 | PILOT4 |
|---------------------------|-------------------|-------------------------------|-----------------|-----------------|
| Date of pilot | 09/10/2013 | 09/10/2013 | 11/10/2013 | 14/10/2013 |
| Stereo vision (out of 10) | - | 10 | - | 10 |
| Age range | 26-35 | 26-35 | 26-35 | 26-35 |
| Gender | Male | Male | Female | Male |
| Lidar expertise | Highly proficient | Proficient | Capable | None |
| Native language | Spanish | English | English | English |
| Occupation | Geography staff | Other: GIS production manager | Geography staff | Geography staff |

3.3.3 Evaluating suitability

Aspects of the pilot study helped to endorse development decisions or inform amendments to the original experiment. These are described below.

Data suitability

Pilot participants were shown the four datasets that are detailed in section 3.2.1 and Appendix C, to test whether (self-ranked) lidar expertise would limit the understanding of the point cloud structure and/or tasks for the main experiment. Prior to the pilot, it was envisaged that the largest AOI (Scene D) may be too difficult to interpret for participants. For the pilot participants who were less familiar with point cloud data (PILOT3 & 4), Scene D did appear to be more difficult to interpret.

Task suitability - measurement

Pilot participants were presented with the areas of interest that were intended to be shown to the main experiment participants (refer to Appendix C – scenes A and B). The task set-up and approach were the same as the measurement task described in section 3.4.3.

The volunteers had varying experience with gamepad controllers (data unavailable), and, during the pilot, the volunteers were given time to practise with the controller. During the task, they were asked to carry out measurements for two point clouds, (1) the longest diameter of a canopy of trees, and (2) the length of a building roof edge. Pilot participants found that they knew which points they wanted to select, but it took time to navigate to the point. Despite different experience with controllers, it was felt that the sensitivity of the controller was too high and meant that movement was not as expected.

Task suitability - interpretation

Volunteers were shown the interpretation scenes (C & D, see figures in Appendix C). Before carrying out the interpretation task, the participants were shown the feature from above, *i.e.* a bird's eye view. The 3D scene was divided

into 8 cells, or segments, and the participant was asked to describe what features they believed to be present in each cell. Although this provided context for the volunteer, it is not the primary objective of the interpretation task and the pilot highlighted that it could potentially use up a significant portion of experiment time.

Like the main experiment, the participants were asked to navigate to and describe 5 flashing points in Scenes C and D. The volunteers had varying experience with gamepad controllers (data not available) and were asked to select and identify the points in 3 minutes. The pilot volunteers reportedly felt pressurised and panicked because of the timer, and they appeared nervous and agitated. Furthermore, when asked to select the POI with the target tool before verbally identifying the point, the participants expressed that they felt pre-occupied with the selection of the point. As a result, the actual task of verbally describing the POI (and providing an answer) became a secondary priority.

Data representation

During pilot tests, the four volunteers were informally interviewed about the colouring of points to determine which representation to use for the main experiment. In standard lidar software, the user can alter the colour and size of the points according to personal preference. However, in the trials, the representation of the data must be consistent for every individual, so that the results are comparative.

Ahead of the main experiment, three of the pilot participants (PILOT2, 3, and 4) were questioned about the colouring of the points to determine which colouring to use for the experiment. They were shown three coloured representations of a point cloud, against a black background:

1. one colour (green)
2. greyscale according to height (z value)
3. red/green/blue (RGB hereafter) according to the three automated classification themes of ground, vegetation >2m, building.

The participants were asked to rank the effect of each representation on their understanding of point cloud structure, under different visualisation modes (2D or 3D) where 1 = very effective and 7 = not at all effective. Those scoring 7 would feel that the representation is not at all effective at helping them understand the structure. Figure 3-12 shows the results, which suggest that the 2D method (top row) benefits from this thematic RGB representation - PILOT2 felt that when viewing the data in 2D, the RGB colouring helped with the understanding of the point cloud structure (rated 1/7). Interestingly, for PILOT2, the RGB method in 3D was less effective. Generally (across 2D and 3D methods), pilot participants felt that it was easier to understand the point cloud structure when the points were categorised into three colours representing automated classification.

Although PILOT1 did not contribute to the ratings in Figure 3-12, they commented on the greyscale z colouring, stating that it was, *'telling the viewer which points are higher than the others'*. In contrast, PILOT3 stated that z value did not give much more information. PILOT1 stated that colour by z values is more advisable when trying to understand the point cloud structure. Instead of one colour and RGB, they felt that when all points are one colour, the viewer has to work out which is higher than the others.

The feedback from the 4 pilot participants highlights that the choice of colour could affect a user's ability to understand a visualised dataset. Furthermore, the effectiveness may vary between 2D and 3D displays. These pilot results helped determine the representation of the lidar points during the main experiment.

Effect of point cloud colouring on pilot participants' understanding of point cloud structure

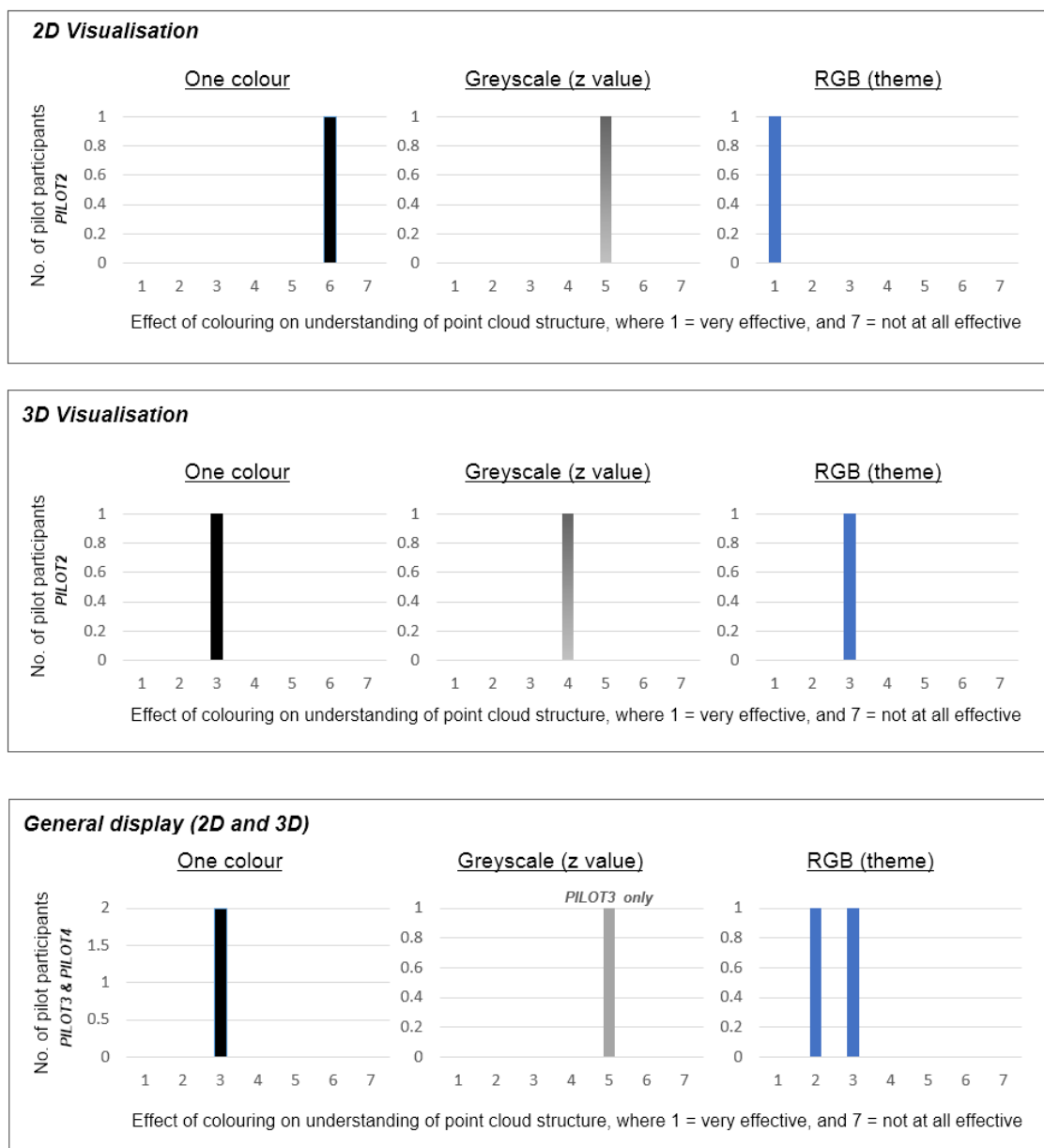


Figure 3-12. Plots showing pilot participants' ranking of different colour representations of lidar point cloud data – one colour (green points with black background), greyscale (by z value), thematic RGB colouring according to classification. X axis ranks each colour representation from 1-7, with 1 being very effective at helping them understand the structure and 7 being not at all effective. Top row and middle row show feedback from PILOT2, bottom row shows results from PILOT3 and PILOT4 regarding general display (i.e. regardless of 2D or 3D method). Note different y axis in one colour general display.

3.3.4 Amendments to the experiment

The findings of the pilot study guided the main experiment with regard to lidar data, streamlining of the experiment, and visualisation system (where point colouring is configured). Appendix D lists the amendments that were required following the pilot study.

Lidar data

All four AOIs that were presented were kept for the main experiment. For the interpretation task, some POIs were moved from the ground surface to more structurally complex places, to create more of a challenge to the participants.

Streamlining the experiment

The flow of the experiment was streamlined following observations of the pilot participants, who were distracted by the alternative tests in the Randot™ Stereo Test (Stereo Optical Co., 2009). These supplementary vision tests were obscured with cardboard (as shown in Figure 3-10). Participants felt that the grid exercise during the interpretation exercise helped with understanding/orientation during closer inspection of the points. As a result, this talk-aloud overview of the scene was kept to help the volunteers gain context from the whole scene, however, the grids cells were reduced to 6 to reduce discussion time. After the reaction of pilot volunteers to time constraints, there was no restrictive time span. This is similar to Van Coillie *et al.* (2014), who carried out a remote sensing operator study with no pressure to complete each task within a certain period of time. The feedback questions were presented to the pilot volunteers in writing, but, for the main experiment, these questions were displayed on screen so that further discussion could be generated and recorded on the Dictaphone and/or on the spreadsheet.

Visualisation system

The pilot study suggested that colouring of points could have an impact on participant understanding of the data structure. As a result, an additional colour

scheme written into the code so that the main experiment first displayed the interpretation task points in one colour, then re-coloured the points according to classification (ground, building, vegetation >2m). Owing to time constraints, this re-coloured interpretation task is not analysed, but it was carried out by the volunteers during the experiment. Participants carried out the interpretation task with the points as one colour, before revisiting (and perhaps revising) their answers while viewing the classified data. Although re-interpretation of the same data inevitable introduces the learning effect, analysis should offer insight into the effect of data representation on the participants' understanding.

Additional contextual information, including as automated classification, is likely to have an effect on the outcome of the visualisation results, however, the main experiment uses one colour for all of the points so that any confounding effects on the 2D and 3D techniques can be separated out. Colouring by greyscale and classification were discounted in order to separate out and reduce the number of conditions that were to be tested. A drawback to reducing the number of confounding factors is that the colouring of the points could affect any disparity between 2D and 3D visualisation methods. Ultimately, the point clouds were coloured one hue, dark blue (RGB = 21, 120, 180), against a light green (RGB = 178, 223, 138) background. The specific colouring was identified using ColorBrewer 2.0 (<http://colorbrewer2.org/>, accessed 08-08-13) to ensure colourblind participants were not at a disadvantage, as this was not tested for during the experiment.

Regarding the interaction device used, the configuration of the gamepad was adjusted in the code and pilot volunteers were recalled prior to the main experiment to assess the changes. This was tested by navigating within a point cloud and selecting points, before the main experiment.

3.4 Main experiment

Once the visualisation system had been developed to show four AOIs, in either 2D or 3D, the data collection was carried out, Figure 3-13. This involves

recruitment of participants, the physical set-up of the experiment environment, an explanation of the importance of scene order, and a description of the stages of the experiment.

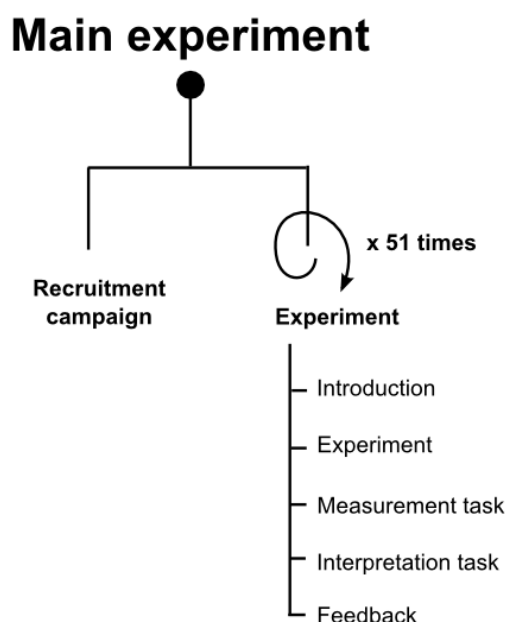


Figure 3-13. Flow diagram highlighting the timeline for the main experiment stage of the methodology

3.4.1 Recruitment campaign

During university term-time, between October 2013 and February 2014, a variety of media were used to attract staff and student volunteers from the University of Leicester, UK, to the experiment. The majority of the adverts (posters, promotional emails, and e-bulletin newsletter advertisements) were university-centric and this is justified by the location of the experiment, which was campus-based. It was assumed that by attracting volunteers who frequent the area, they would be more inclined to take part in the experiment. As a result of the advertising campaign, approximately 80 expressed interest in taking part in the study and bookings were made online, via WeJoinIn sign-up tool (wejoinin.com, last accessed 28-02-14). Prior to the experiment, potential participants were issued, via email or in person, an Information Sheet and Consent Form (Appendix G) the former of which was provided to help them evaluate whether or not they would like to take part.

3.4.2 Experiment set-up

VR theatre

The Virtual Reality (VR) Theatre in the University of Leicester's Geography Department provided the experiment setting, which was set-up as shown in Figure 3-15. The researcher (R) had control of the visualisation system via a wireless mouse and keyboard.

Order of experiment tasks

If a volunteer measures a building via a 2D display, then carries out the same task in 3D, the knowledge gained from their first attempt may influence the second measurement. So that this learning effect is reduced, the participants carry out tasks on the 4 different AOIs in a random order. This random allocation also helps to ensure participant background does not influence the success of the (2D or 3D) approach, after Gigli and Casagli (2011). Figure 3-14 details the order in which the participants viewed the scenes and which method was used. Appendix E lists the methods used by each participant per scene. Between the two tasks, repetition of method (2D or 3D) was not permitted so that each participant could take a break from flickering of the shutter glasses.

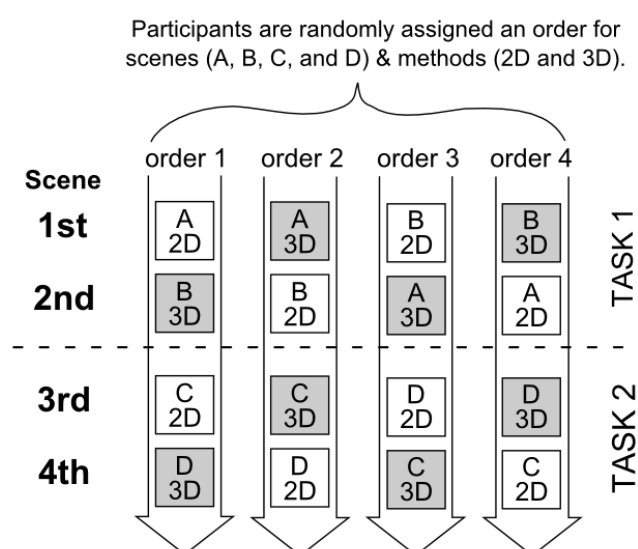


Figure 3-14. Different orders, relating to the combinations of scenes (A, B, C, and D) and visualisation methods (2D or 3D) that were randomly assigned to participants.

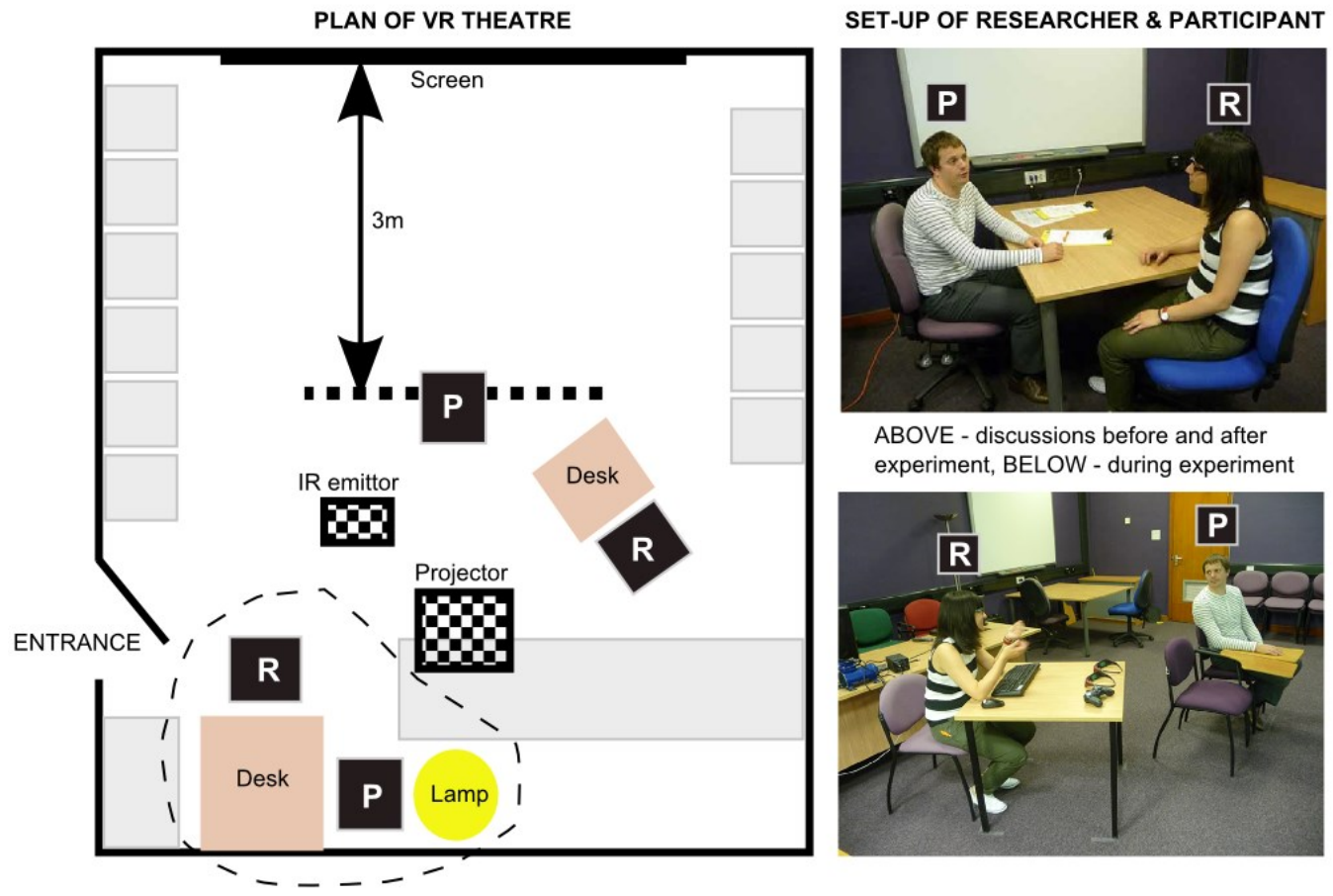


Figure 3-15. Schema (left) showing the layout of the Virtual Reality Theatre within the Geography Dept at the University of Leicester, which was used during trials. The room is approximately 6m x 7m. The chequerboard elements represent the 3D stereo hardware, which is mounted on the ceiling. All other grey elements are furniture, desks or chairs, that were present, but not required. Dashed line indicates area for introductory discussions and feedback, prior to the experiment. Top photo shows the set-up, where R = researcher's position and P = participant.

3.4.3 Experiment stages

A description of the stages of the experiment are described, with the symbol ► denoting the generation of data by the participant. A detailed flow diagram of the experiment, inclusive of measurement and interpretation tasks, is presented in Appendix F and the interviewer script is included in Appendix H. One interviewer carried out all of the experiments, which help avoid the evaluator effects of multiple moderators (Hertzum and Jacobsen, 2001).

STAGE 1 - introduction

The participant and interviewer begin the experiment at the rear of the VR room (area outlined by dashed line in Figure 3-15) to review the Information Sheet and Consent Form, before signing. Participants are also encouraged to ask any questions that they may have regarding the experiment.

Vision test

Each volunteer was asked whether they have any issues with their vision and their stereoacuity was measured (Figure 3-16 - Screening) using the Randot™ stereo test (Stereo Optical Co., 2009), pictured in detailed in Figure 3-10. It was explained that those who do not have a score of 6/10 are unable to carry out the full trial.

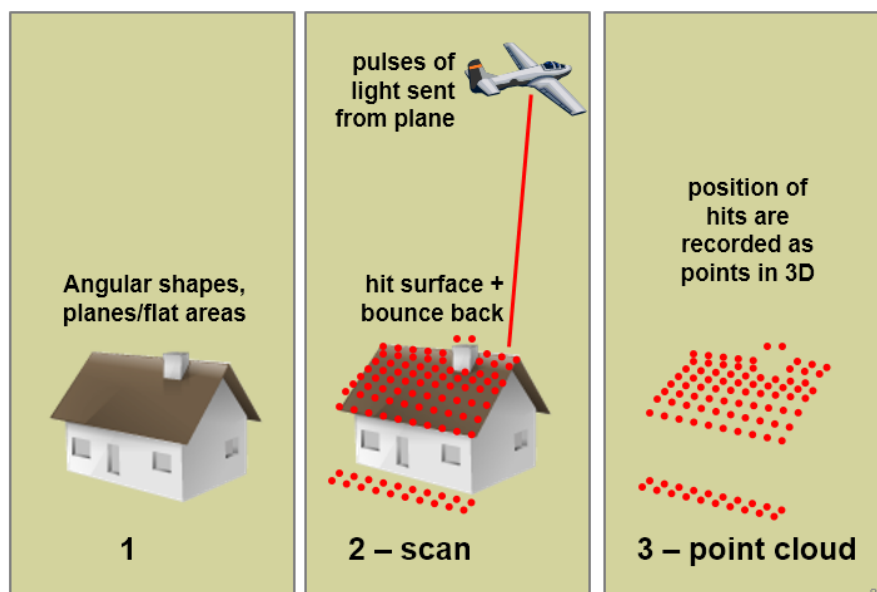
Familiarity with lidar

The interviewer questions the volunteer's familiarity with lidar and laser-scanning. The concept of lidar was then explained to each participant, using the images in Figure 3-17 (displayed on the projector screen), to make sure that all participants have the same baseline knowledge before undertaking the tasks. Participants were told that, in general, point clouds derived from man-made features were more planar and regular in shape, and there could be shadow in the point clouds. Natural features were described as being more irregular or messier, and any gaps between leaves and branches might allow the laser to penetrate through to the ground.



Figure 3-16. Experiment flow - the participants who passed the stereo screening test and continued to carry out the main experiment.

Point cloud in 3 steps: man-made features



Point cloud in 3 steps: natural features

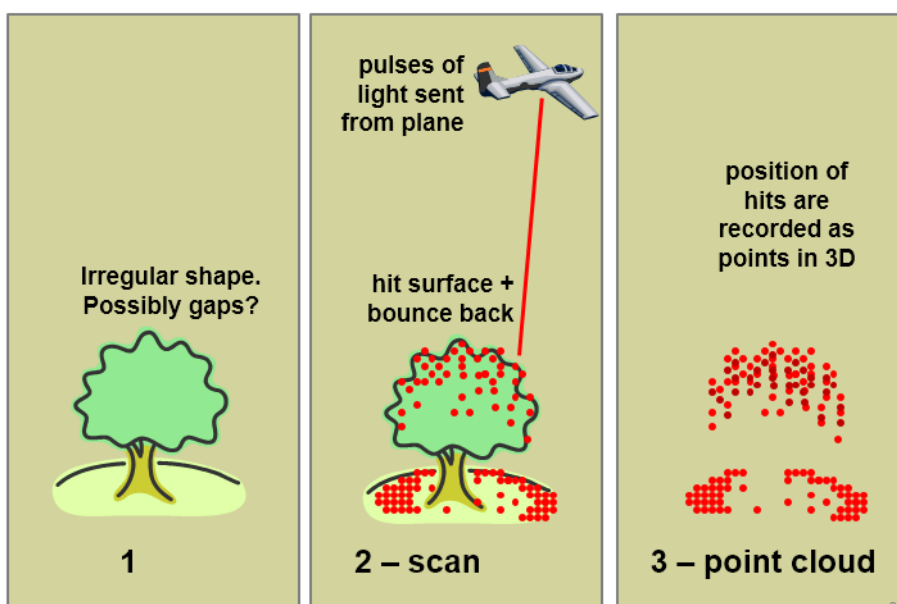


Figure 3-17. Two slides presented to the participants to explain the concept of lidar data acquisition for man-made features (top) and natural features (bottom). Images author's own, produced using Microsoft PowerPoint (Microsoft Corporation, 2010).

STAGE 2 – main experiment

The participant was relocated and seated 3m from the visualisation screen (Figure 3-15, right), at this stage, the distance between the pupils of the participants was measured (Figure 3-16, top row). This was then inputted to the system, by the interviewer. The participant undertook gamepad training so that they are were to navigate through the point clouds and explore their structure at will. The Saitek controller (Saitek, 2007) is introduced, alongside visual instructions (Figure 3-18) for (a) looking and (b) moving. The participant is given time to practise by exploring their first scene (either A or B) until they feel confident to continue. On average, this took 4 minutes.

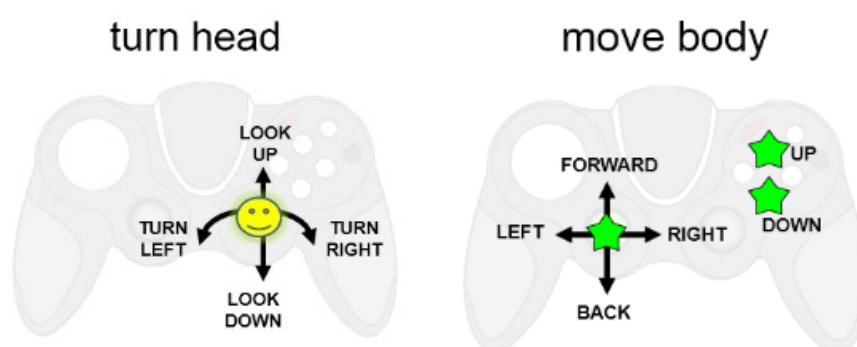


Figure 3-18. Summary of participant guidelines for gamepad training. Left analogue stick = movement, right analogue stick = head turn/look.

During task 1, which is further described in Chapter 5, volunteers were shown instructions to measure Scene A and B (Appendix C) using either 2D or 3D visualisation methods. The exposure to the different point clouds in Scene A and B, and their reference imagery, provided visual object recognition training in preparation for the interpretation task (task 2). Participant instructions for task 2 are detailed in the Chapter 6. These were carried out via either the 2D or 3D method, using Scenes C and D (Appendix C) without reference imagery. Throughout the course of the two tasks, participants were shown the four AOIs, firstly from plan-view, then North-, East-, South-, and East-facing views. Ahead of their own navigation of each scene, they were positioned at the North-facing view. Images of these views, for Scenes A to D, all of which are displayed in Appendix C.

STAGE 3 - Feedback

Once the experiment tasks were complete, the participants were asked to reflect on the tasks, during the feedback session (Appendix H). The user responses were recorded directly into a spreadsheet by the interviewer, using Excel® (Microsoft Corporation, 2012a), as well as via the Dictaphone. A written survey was also carried out (Appendix I), at the end of the experiment to gather information about user background.

3.4.4 Data Outputs

The outputs of the experiment are summarised in Table 3-9 (each denoted by symbol ►) and the underlined text highlights the datasets that are carried through for further processing and analysis. Appendix F pictorially (also using ► symbols) shows the stages of the experiment from which they are derived.

Table 3-9. List of data outputs from relevant phases of the experiment. Underlined text denotes data used in results. Stereo data are only used to determine participant inclusion and visualisation set-up.

| Phases of Experiment | Quantitative Data | Qualitative Data |
|------------------------|---|--|
| Consent | ► Signed hardcopy forms | |
| Stereo test & set-up | ► Stereoacuity score used for screening. ► IPD noted, inputted into system. | |
| Task 1: Measurement | ► <u>.txt file</u> - for Scenes A and B: on button press of first and last points, records selected point's xyz coordinates and time pressed. | ► <u>audio</u> - associated comments, if given. |
| Task 2: Interpretation | | ► <u>audio</u> - verbal classification answers for Scenes C and D ► <u>audio</u> - associated comments, if given. |
| Feedback Questions | ► <u>.xls</u> - answers inputted into Excel spreadsheet | ► <u>audio</u> - answers recorded |

3.5 Data Processing & Analysis

In this section, the data processing and analysis of the output experiment data is outlined generally, with reference to the required software, Table 3-10.

Further to the general methodological development that is described in this chapter (pictured in Figure 3-19), methods that are specific to the measurement and interpretation tasks are detailed in Chapters 5 and 6.

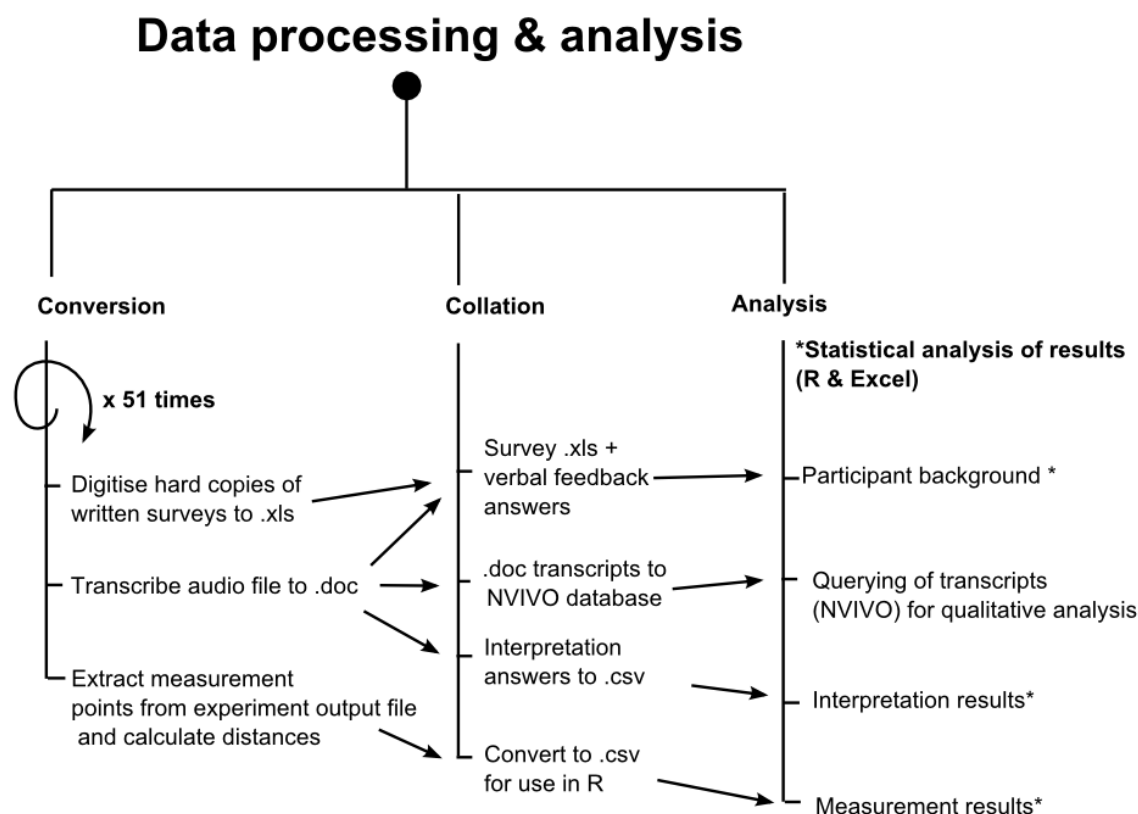


Figure 3-19. Diagram of the data processing stage – data conversion, collation and analysis. The analysis methods used for each task are detailed in chapters 5 and 6.

3.5.1 Experiment data conversion and collation

The audio recordings amounted to roughly 90 hours and ExpressScribe[®] playback software (NCH Software, 2013) was used to facilitate the manual transcription of .wma files, which took approximately 4 months to complete

singlehandedly. The transcript of each participant was written into a separate Word® (Microsoft Corporation, 2012b) document, which were then imported into NVIVO® (QSR International, 2013), a qualitative data analysis software package. Pre-determined heading-styles of the imported Word transcriptions (Microsoft Corporation, 2012b), which split each experiment into consistent sections, meant that the NVIVO software could automatically divide the 51 transcriptions for easy analysis. A database of all of the participants' audio transcriptions was contained within an NVIVO project (QSR International, 2013), from which themes were highlighted and allocated to nodes.

Table 3-10. List of software used for processing data generated by experiment.

| Type of Data | Software | Function |
|--------------|--|---|
| Qualitative | • ExpressScribe (NCH Software, 2013) | • Playing audiofiles |
| | • Word (Microsoft Corporation 2012) | • Writing of transcript |
| | • NVIVO (QSR International, 2013) | • Collation of transcripts |
| | • Excel (Microsoft Corporation, 2012a) | • Quantification of feedback |
| Quantitative | • Excel (Microsoft Corporation, 2012a) | • Collation of verbal feedback • Digitising of written surveys |

The transcripts relating to the interpretation task were copied into an Excel file, which also listed some of the quantitative data. This provided a database from which variables could be analysed. The output text files contained data generated during the measurement task. The structure of each of the .txt files was not written in a straight-forward arrangement, so the coordinates of points selected during the measurement task were manually extracted from each output file, from each participant, using Excel (Microsoft Corporation, 2012a).

These coordinate values were used in the measurement analysis. As stated previously, the task-specific analysis steps, such as statistical tests carried out using open-source R package (R Core Team, 2014), are detailed within the relevant task chapters. Unlike age-ranges and Likert scales, which were pre-determined in advance of the study, elements of the results required classification after results had been generated. In particular, the interpretation results required classification.

3.5.2 Data Analysis

The methods used to analyse the measurement and interpretation results are detailed within the chapters 5 and 6. The participants' responses to the written surveys, relating to participant background, were collated in Excel (Microsoft Corporation, 2012a). The results are reported within Chapter 4 as graphical plots, illustrating the characteristics of the sample population. The feedback questions that had been recorded in Excel at the end of the experiment, which were also recorded on the dictaphone, were added to the database. The audio data were analysed to provide a narrative to the quantitative data outputs.

Descriptive statistical analysis

Throughout the results chapters, data are typically displayed as a histogram or as boxplots. For reference, Figure 3-20 demonstrates a normal frequency curve, or histogram. In Figure 3-20, the dark regions below 2.5% and above 97.5% represent the 0.05 or 95% confidence level, against which statistical tests are normally measured. Figure 3-21 is presented as a reminder of the elements of a boxplot and the statistics it represents.

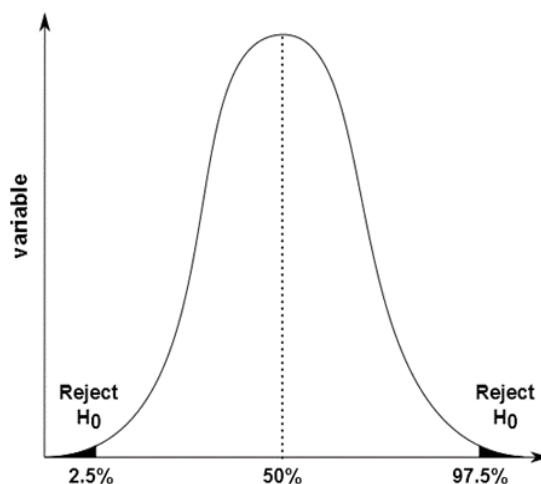


Figure 3-20. Schema showing a normal frequency distribution (in percent) of a sample against a variable. Shaded portions under the bell-shaped curve show the 5% critical region at which the null hypothesis (H_0) is rejected. Image author's own.

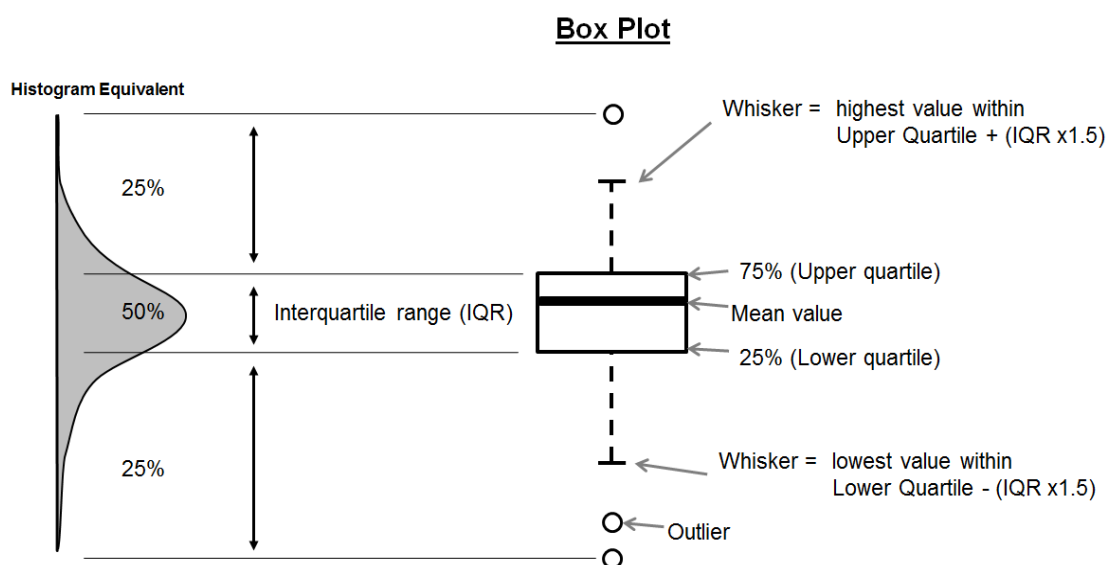


Figure 3-21. Schema to show elements of a box plot, in relation to the equivalent histogram. Absence of the IQR box indicates $\geq 50\%$ of the sample population exhibited the same value. Image author's own.

3.6 Summary

The methods described in this chapter are summarised pictorially in Figure 3-22.

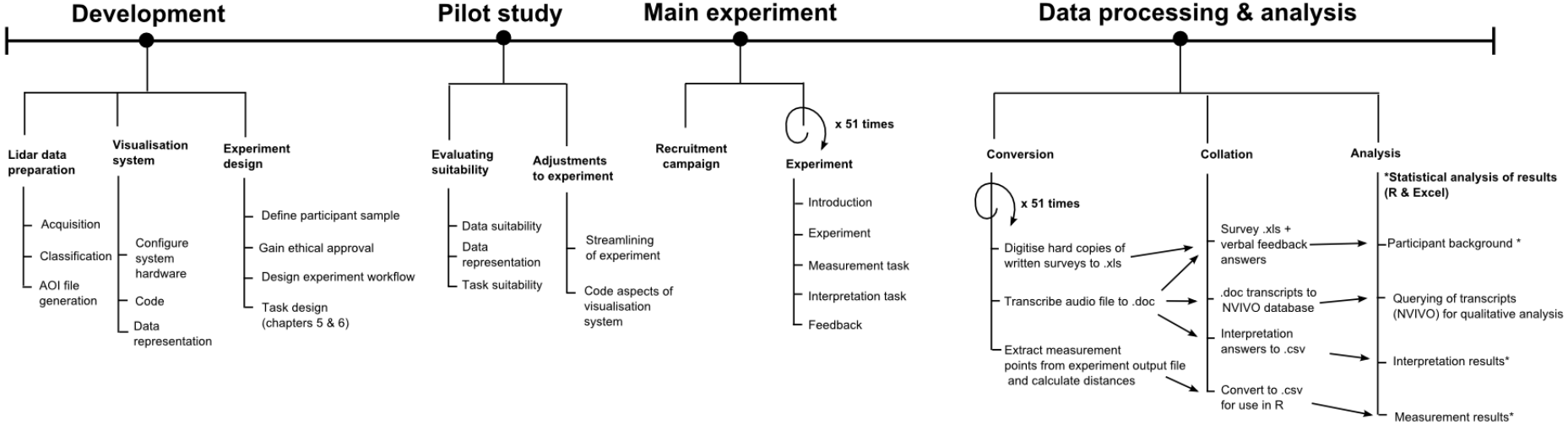


Figure 3-22. A summary of the method - nodes relate to the different stages of project development.

4.

Participant background

4. Participant background

4.1 Introduction

This study was open to anyone aged 16 and above with any level of lidar expertise, in order to attract a number of volunteers to take part in the experiment. A total of 51 volunteers took part in the trials and their general background characteristics were measured and are reported in this chapter.

4.2 Method

4.2.1 Demographic data

Data collection for the participant profiling took part (a) during the introduction of the experiment and (b) through the written survey, at the end of the experiment. This meant that prior to the experiment, certain elements of the data collection were categorised, *e.g.* age range. The original questions can be referred to in Appendix H and I. Data were visualised using R (R Core Team, 2014) or Excel (Microsoft Corporation, 2012a) and are displayed in sections 4.3 – 4.5.

4.2.2 Ranking participant lidar expertise

Unlike the study's demographic analysis in Chapter 3, which was categorised prior to the experiment (*e.g.* age range), some aspects of the data collection required ranking after the experiment. Table 4-1 shows the categories used to rank lidar knowledge and experience into novice, informed, and expert levels. This method was developed in response to the issues associated with categorising expertise, and is further described in the text that follows Table 4-1. In the table, the notation of P followed by a number indicates the participant from whom the quote is taken, *e.g.* P42. This format is used hereafter when making direct reference to a particular participant.

Table 4-1. Ranking of participants for lidar/laser-scanning (a) knowledge and (b) experience, based on transcriptions. Where 0 = novice, 1 = informed, 2 = expert.

| Rank | Level | (a) Knowledge example | (b) Experience example |
|------|----------|---|--|
| 0 | Novice | P36 To be honest with you, I studied this course in my masters, but... and I KNOW about it- it's one of the technique in remote sensing, capture the land's action, but... in details, I don't remember anything! | P34 <Seen colleagues using it on screen, in passing > |
| 1 | Informed | P43 Um, I think it's something to do with lasers, is it measuring, um, sort of height of stuff. I don't really know. | P35 From a colleague who is using lidar data. My colleague uses lidar data for estimation, which is, er, main part of my work, so I do assist him in doing one or 2 things with lidar data. |
| 2 | Expert | P28 Okay, well, it's usually used from an airborne instrument, like, from aeroplane or even helicopter. So, it's a sensor or scanning device that actively sends out a laser and retrieves it back again. And by that you can gauge, like the height of vegetation, so it has a first return and last return, so you will get the, for example, the height of the canopy. PLUS from the first return and the last return, would be actually the digital elevation mode, that you can derive from that. Yeah and you can do airborne and terrestrial lidar scanning. | P03 Yeah, been working with a satellite... It's called ICESAT GLAS... waveform to estimate mean canopy height. It's the software from ICESAT GLAS, they have, I don't remember now the name. And I processed just first and the rest were processed from another organisation. Yes is from the waveform, just using the waveform, like different parts of the waveform, you estimate the canopy height just using a linear model, polynomial model. I have to take from the software, you get the values and then estimate the canopy height afterwards. |

Geovisualisation studies commonly split participants into novice (or naïve) users versus experts (Hegarty *et al.*, 2009) and during the first stage of the main experiment (described in section 3.4.3) participants were asked to rank their level of lidar expertise on a scale ranging from 'very proficient' to 'none'. If the participant expressed their level expertise as 'none', no further questions

were asked regarding lidar expertise or familiarity. If any other answer were given (novice/beginner, capable, proficient user, or very proficient user), the interviewer prompted the participants for further comment regarding their past experiences (including software used, analysis performed, *etc.*).

Roth (2015) highlights the complexities of measuring performance against expertise; expertise is not discretely categorised and the boundaries between categories are uncertain. For these reasons, the expertise ranking was re-categorised into knowledge and experience after the experiment, based on the participant transcripts. The participants were then split into 3 groups, similarly to Slocum *et al.* (2001), who categorised novices, informed, and domain experts. Table 4-1 gives an example of the categorisation of participants' knowledge and experience, based on their past experience comments. This ranking was applied to participant knowledge and participant experience with lidar/laser-scanning for all participants who took part in the main experiment.

4.3 Characteristics of overall sample population

The results in section 4.3 give an overview of the background of all 51 volunteers, inclusive of those who scored <6/10 in the stereo test. Section 4.5 describes the characteristics of the 46 participants who took part in the main experiment.

4.3.1 Demographics

Most participants were male (63%), Figure 4-1, and almost half of the volunteers were 25 years old or younger, Figure 4-2. The lower ages could be attributed to the high proportion of students that made up the sample population (see Figure 4-3).

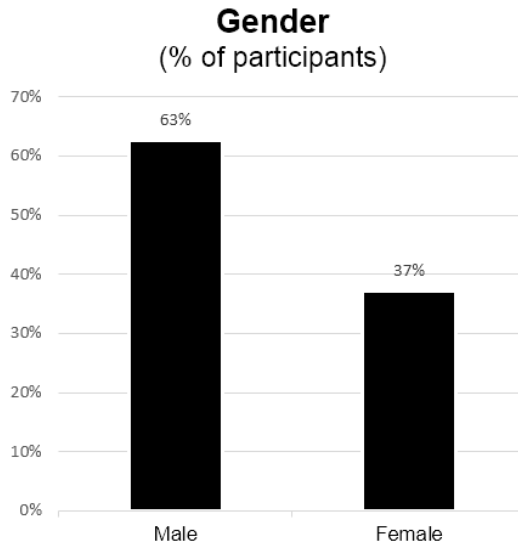


Figure 4-1. Gender split of volunteers.

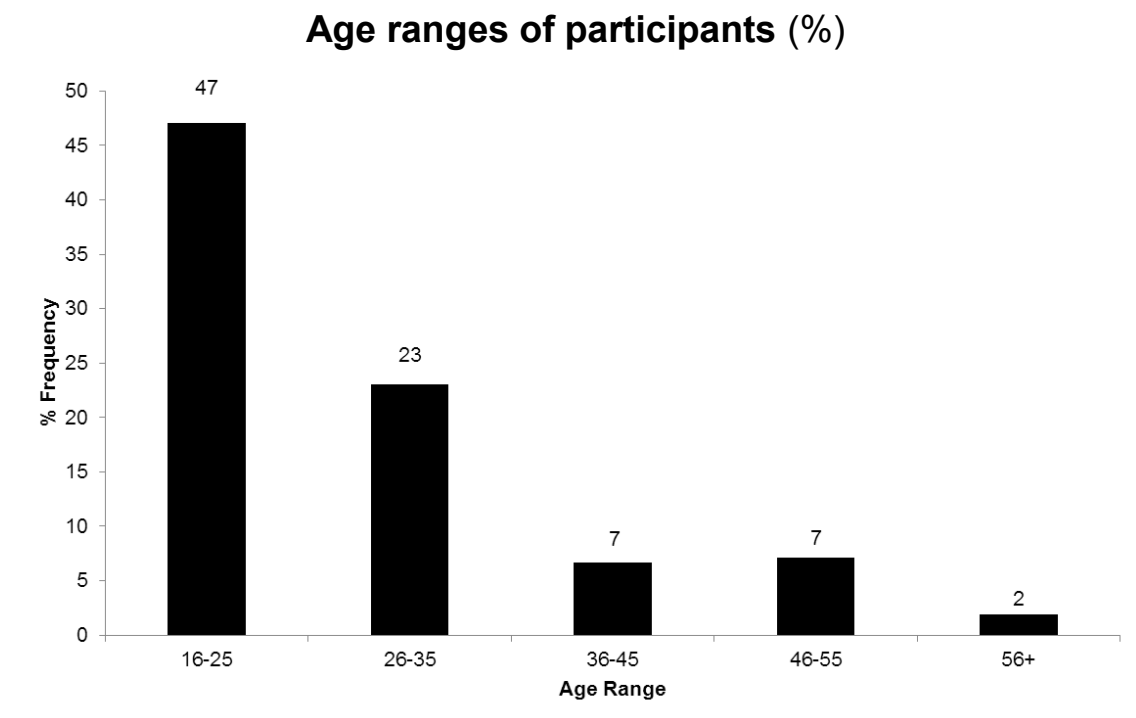


Figure 4-2. Distribution of age ranges of volunteers.

The range of subjects studied spanned Geography, Archaeology, Geology, Cognitive Science, and Business Studies. Those participants who were unrelated to the university, *i.e.* neither student nor staff, were unemployed, housewife, self-employed, care-worker. Kastens and Ishikawa (2006) found that expert geoscientists performed better through repetition of task and practice (p.55). However, in this study, for each task, the participants were randomly allocated to 2D and 3D groups. The study design mitigated against bias expertise that could affect the results. The background of the participants, whether they study or work in the geospatial field, is not a precursor to spatial ability; a participant who has high spatial ability may not be in the geography profession. However, the knowledge and experience of participants who took part in the main experiment are shown in section 4.5.

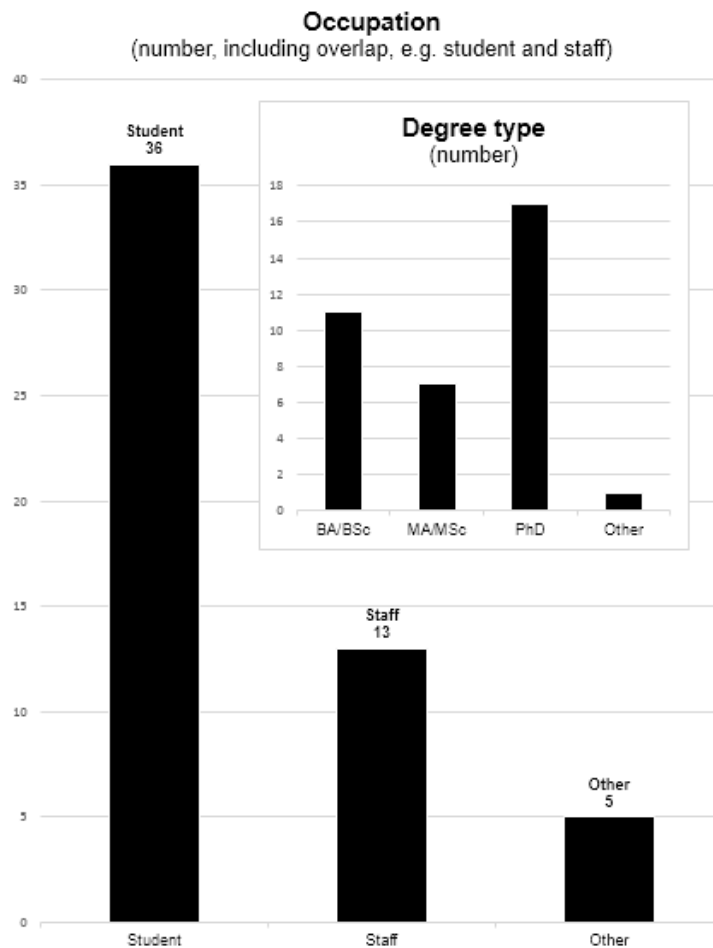


Figure 4-3. Graphs describing the occupation of the sample population. Inset graph shows the level of degrees being pursued by the student category.

Although two-thirds of the participants were native to the UK (Figure 4-4), some native English speakers were less expressive than non-native speakers. It was observed that the different level of English language skills (Figure 4-5) did not necessarily differentiate between those who were more/less communicative during the trial.

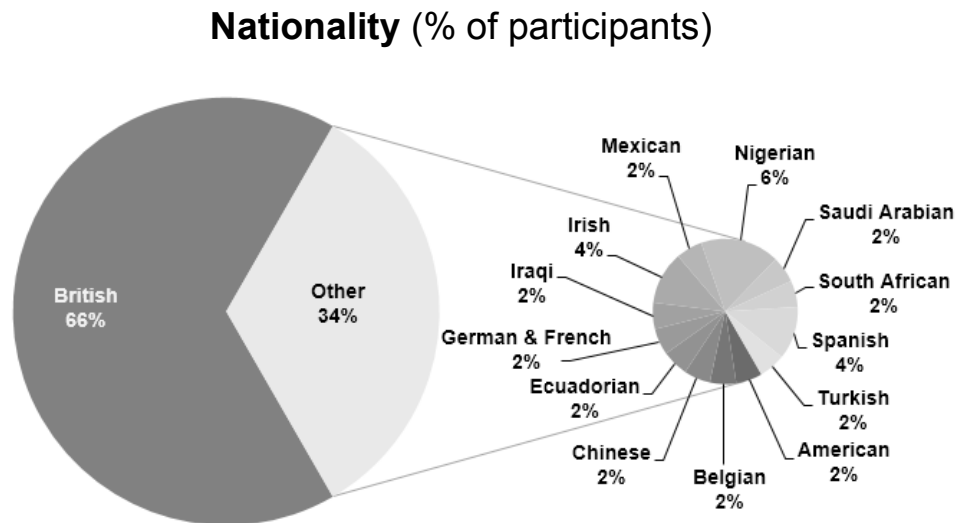


Figure 4-4. Nationality of all participants.

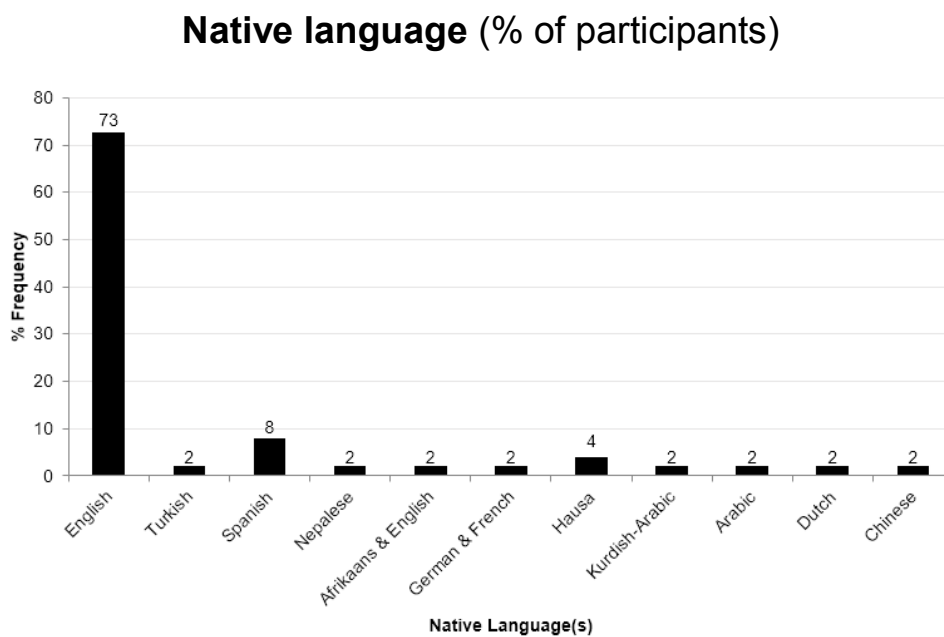


Figure 4-5. Native language of all participants.

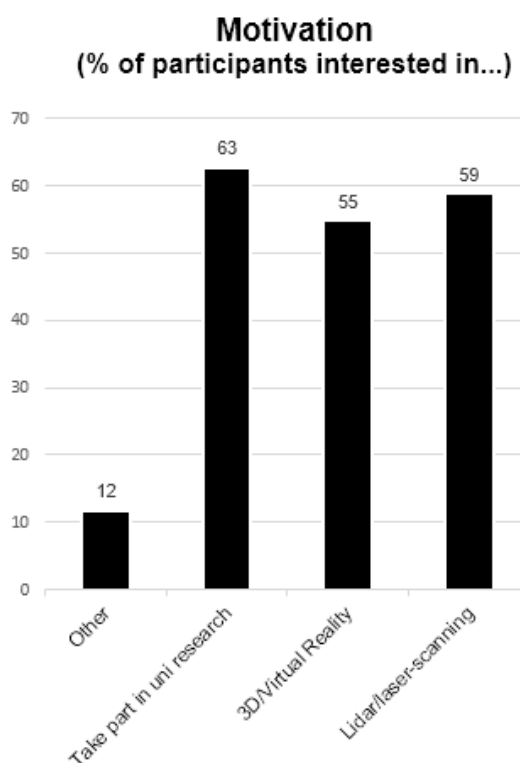


Figure 4-6. Participant motivation to take part in the study.

4.4 Screening of participants

Before the visualisation tests, participants were screened to test their level of stereovision and the result of the vision test determined whether an individual was allowed to continue with the full experiment. Volunteers were excluded from the study if they could not correctly answer 6 or more of the 10 stereo images in the Randot™ test (Stereo Optical Co., 2009). This is equivalent to a stereoacuity measurement of 50 seconds per arc (at 16 inches viewer's eyes).

4.4.1 Stereoacuity results

Of the total sample population, of which there were 32 males, 19 females, five of the volunteers (all male) did not meet the stereoacuity requirements, see Table 4-2 and Figure 4-7.

Table 4-2. Scale of score card results for the Randot™ test (Stereo Optical Co., 2009). Number and percentage of participants (male and female). A score of 6/10 or above lead to the volunteer's inclusion in the full visualisation experiment.

| Ability to see 3D objects in stereo | Randot test score (out of 10) | Stereoaucuity (secs per arc) | No. of participants | | % of participants (n=51) | |
|-------------------------------------|-------------------------------|------------------------------|---------------------|------|--------------------------|-------|
| | | | Female | Male | Female | Male |
| none/stereoblind | 0 | None | 0 | 4 | 0.00 | 7.84 |
| LOW | 1 | 400 | 0 | 0 | 0.00 | 0.00 |
| | 2 | 200 | 0 | 0 | 0.00 | 0.00 |
| | 3 | 140 | 0 | 0 | 0.00 | 0.00 |
| | 4 | 100 | 0 | 0 | 0.00 | 0.00 |
| | 5 | 70 | 0 | 1 | 0.00 | 1.96 |
| HIGH | 6 | 50 | 2 | 1 | 3.92 | 1.96 |
| | 7 | 40 | 3 | 4 | 5.88 | 7.84 |
| | 8 | 30 | 1 | 4 | 1.96 | 7.84 |
| | 9 | 25 | 9 | 5 | 17.65 | 9.80 |
| | 10 | 20 | 4 | 13 | 7.84 | 25.49 |

Stereoaucuity results (% of total number of participants)

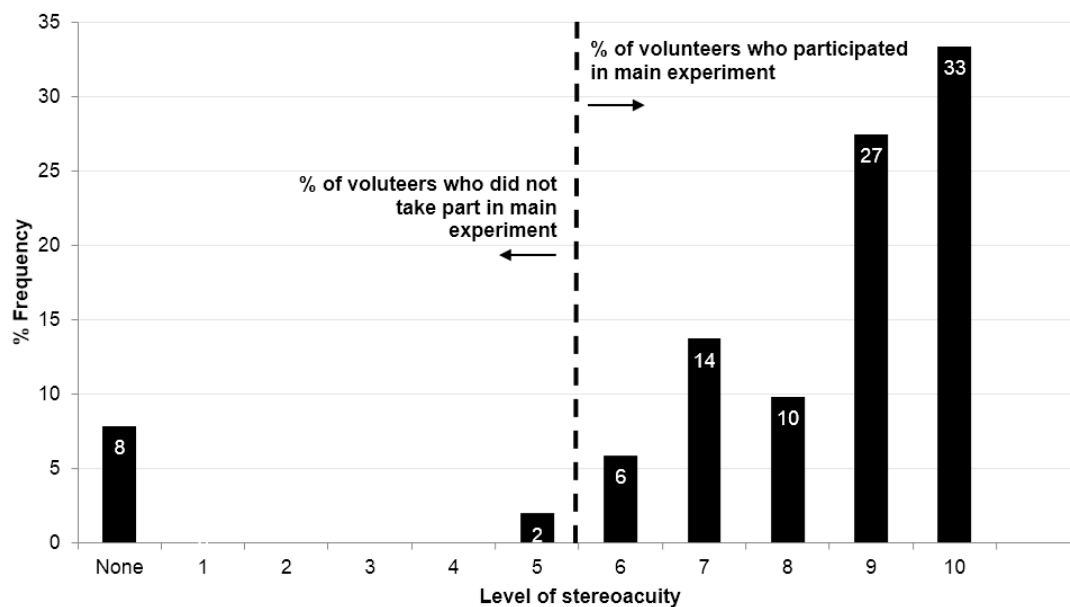


Figure 4-7. Bar graph showing the levels of stereoacuity reached by the general sample of participants (n= 51). Results of 6/10 or above meant that participants were not allowed to take part in the full experiment. 6/10 or higher is the equivalent of 50 seconds of arc or lower.

Two of the volunteers who did not pass the screening test were over the age of 56. Zaroff *et al.* (2003) found that an analysis of variation test showed that stereoacuity is influenced by age and this is thought to be related to cerebral factors. There is a deterioration of stereoacuity with age, particularly over the age of 60. The study did not determine which factors cause the participants' low stereoacuity scores (of 0 and 5), but the screening test caused a reduction in the study's age range to between 16 and 55. The stereo score frequency distribution of the total population are plotted according to gender, Figure 4-8 (left), alongside the sample of participants who passed the stereo test (n=46, right).

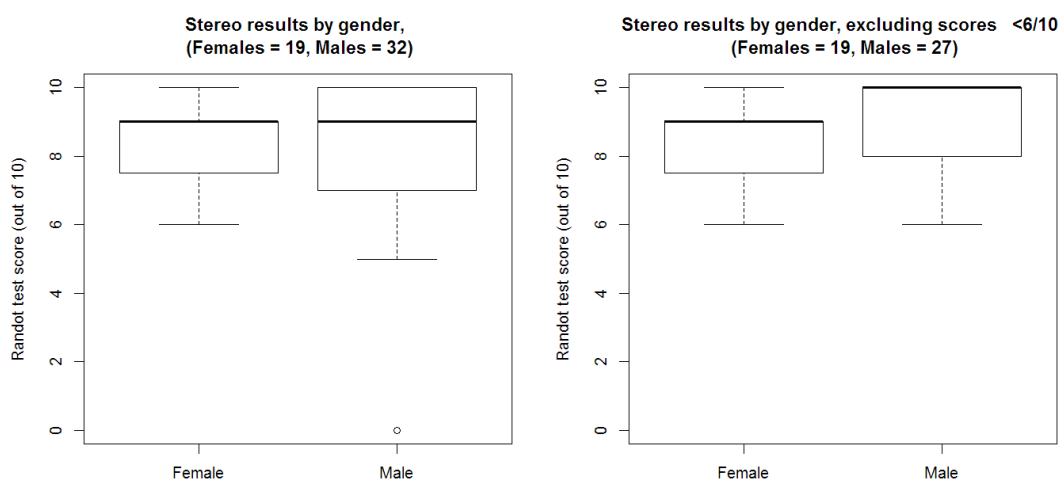


Figure 4-8. Boxplots showing the distribution of stereo scores achieved by females and males who (left) volunteered to take part in the experiment and (right) passed the stereo test and took part in the main experiment. Dark line denotes median.

Bias test - gender

Cognitive psychology literature maintains that stereoacuity is not affected by gender, however, the excluded participants were all male (Table 4-2). Although there are more male volunteers, compared to females, gender bias may exist in the selection process. Statistical tests were run to determine whether there was bias between the stereo scores of the males and females of the entire sample population. First, the Shapiro-Wilk test of normality (Shapiro & Wilk, 1965), using the `shapiro.test()` function in R (R Core Team, 2014), revealed that the distributions of each group are normal (female group: $W = 0.832$, p -value =

0.00346; male group: $W = 0.69903$, $p\text{-value} = 8.407e-07$). A subsequent parametric comparison of the overall male and female group test scores, using the $t.test(x,y)$ function in R (R Core Team, 2014), shows that their results are not significantly different ($p\text{-value} = 0.221$). Therefore, the rejection of participants based on stereoacuity is not biased. Table 4-3 summarises the t -test results of male vs. female participants before and after stereo screening. The increase in the p -value to 0.268 means that the female and male stereo score distributions became more similar after the omission of the 5 males (and the number of participants in each group became more equal).

Table 4-3. Parametric t -test between females and males for all stereoacuity scores and those over than or equal to 6/10.

| | Female n | Male n | p-value |
|------------------------------------|----------|--------|---------|
| All stereo results 0/10 – 10/10 | 19 | 32 | 0.221 |
| Stereo results >>6/10 | 19 | 27 | 0.268 |

Although the ratio of men and women is unequal in this study, *i.e.* 27:19, the individuals were randomly allocated to the 2D or 3D method, for each scene they examined. The self-selection of participants and the subsequent omission of 5 participants from the study, based on stereo test scores, has not resulted in a gender bias.

4.5 Characteristics per trial group

Each participant carried out the experiment in a certain order, as explained in section 3.4.2, to reduce the learning effect between different scenes and methods. This split in participants into 2D and 3D subgroups for each scene, resulting in distribution of demographic characteristics, stereoacuity, and lidar familiarity.

4.5.1 Demographics per 2D/3D group

The distribution of ages per scene are shown in Figure 4-9, which is divided into 2D groups (top) and 3D groups (bottom). The gender split for each of the groups is displayed in Figure 4-10. In all scenes, there is a higher proportion of males and the majority of participants are aged 35 and under.

Age distribution of 2D and 3D groups, for each visualised scene

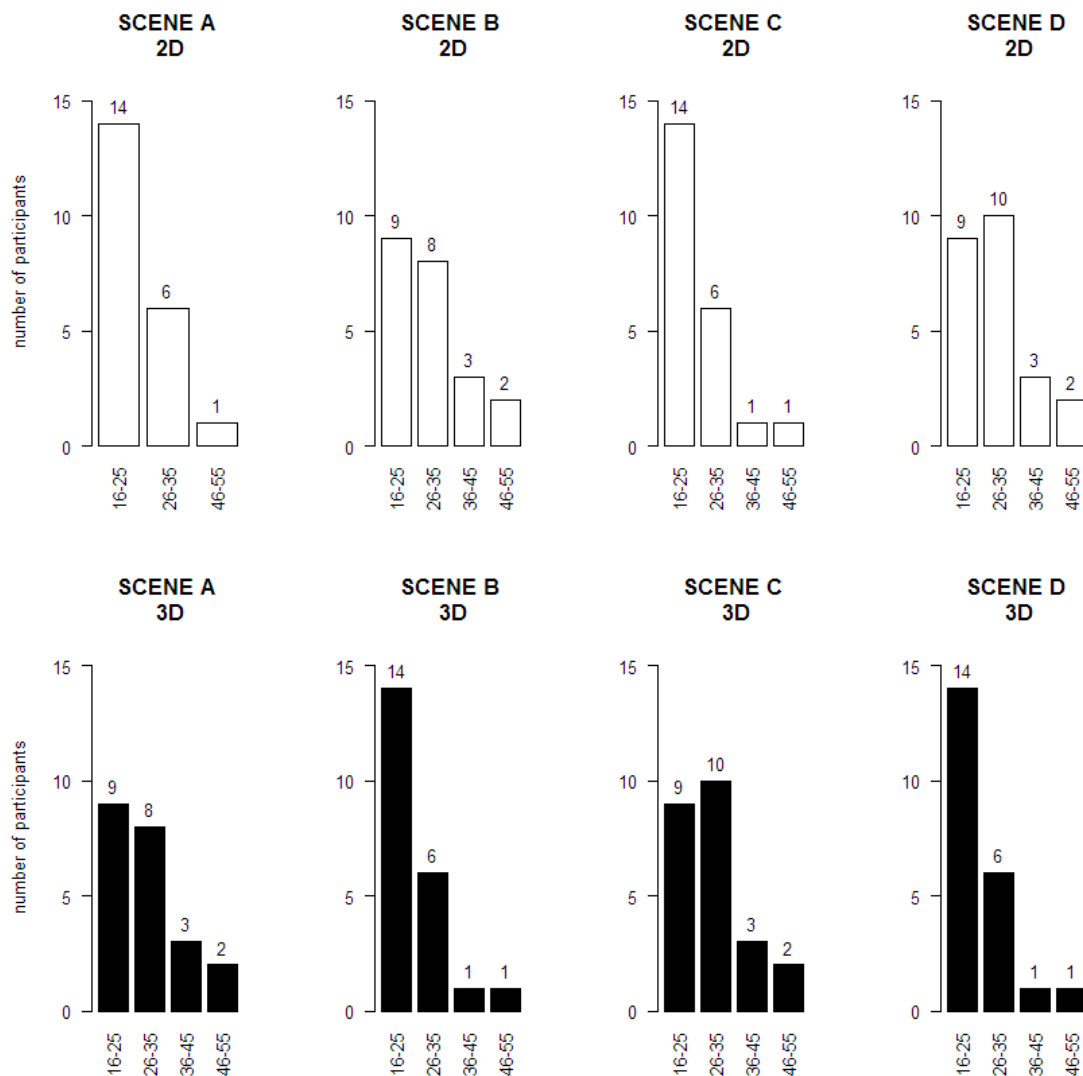


Figure 4-9. Age-ranges of participants for Scenes A to D, for 2D (top) and 3D groups (bottom).

Gender split within 2D and 3D groups, for each visualised scene

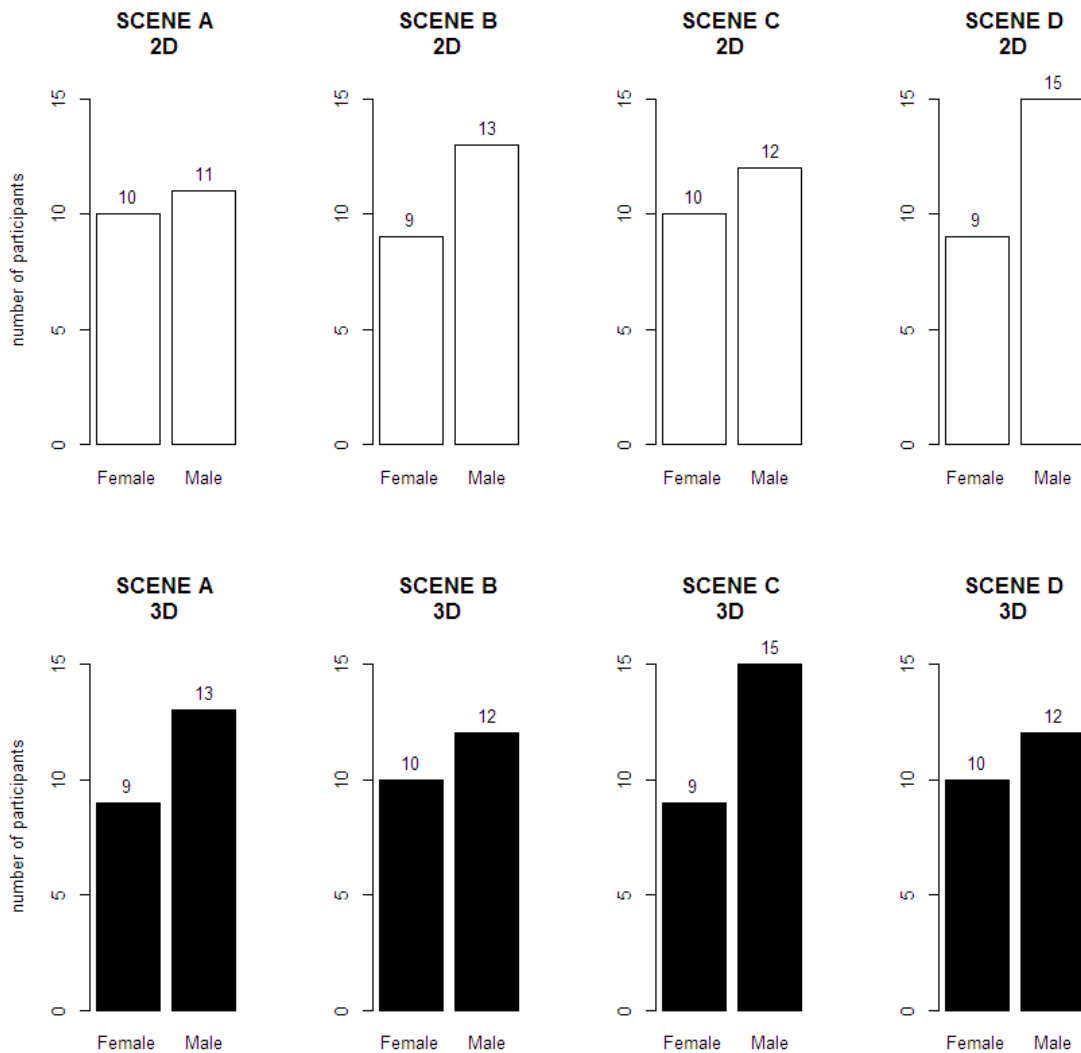


Figure 4-10. Barcharts showing gender split of participant groups for Scenes A to D, for 2D (top) and 3D groups (bottom).

Stereoacuity per 2D/3D group

The boxplots in Figure 4-11 divide the participants into 2D and 3D groups, for each of the Scenes that were visualised. During the interpretation task, in which volunteers viewed Scene C and D, there is a difference between the stereoacuity of the 2D and 3D groups for each scene. Scene C has a larger range in the 3D group, whereas Scene D has a smaller range at a higher score. This uneven distribution of stereoacuity may or may not bias the outcome of the

interpretation results. The employment of participants with high stereoacuity during a 3D test could have a favourable outcome towards that method. Had the allocation of participants to 2D and 3D methods been sorted according to stereo test score, a more even spread of stereo test results could have produced a different outcome in the test. However, the screening process of the experiment already truncated the rankings to 6/10 - 10/10; in essence, the study retained those with good-to-excellent stereo vision. The results shown in Figure 4-11 only provide a more detailed categorisation of this prerequisite. Furthermore, logistically, it would have been difficult to allocate participants before knowing the overall distribution of the sample population's stereo results.

Stereoacuity scores for 2D and 3D groups, for each Scene

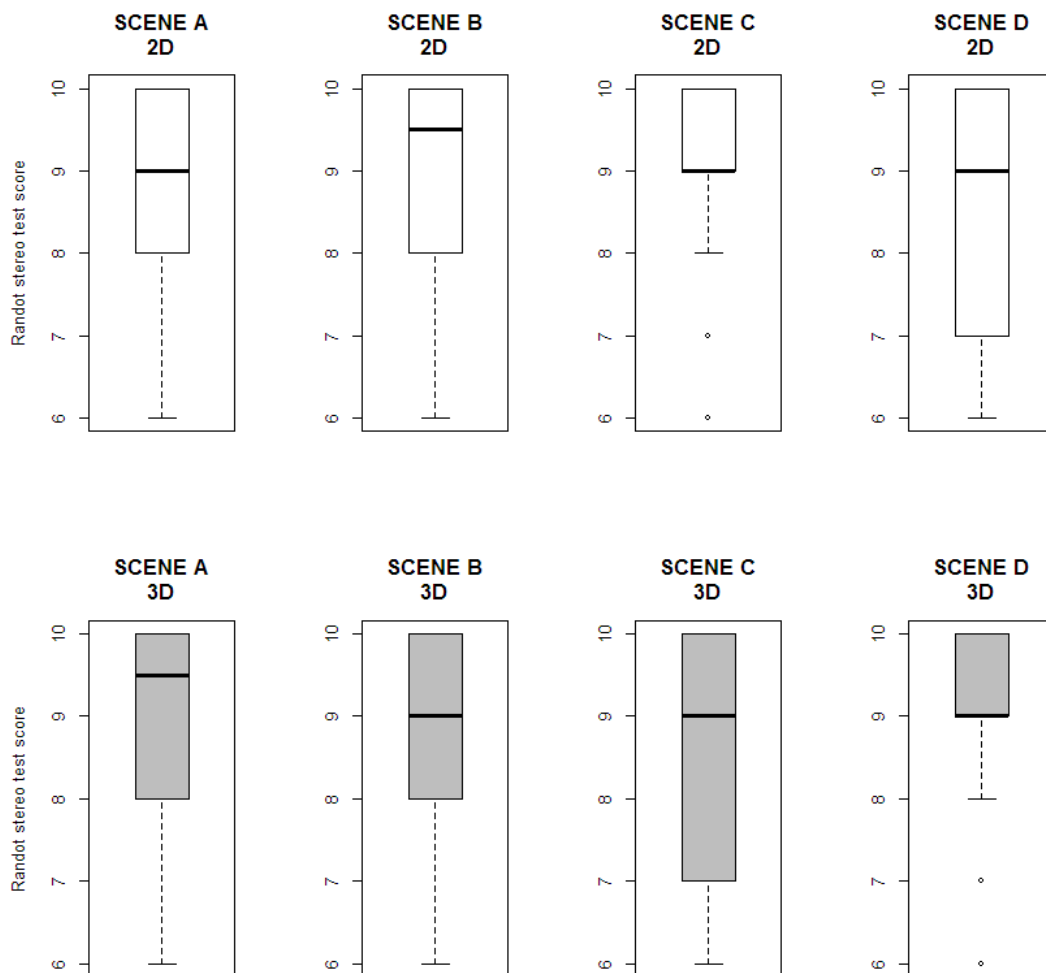
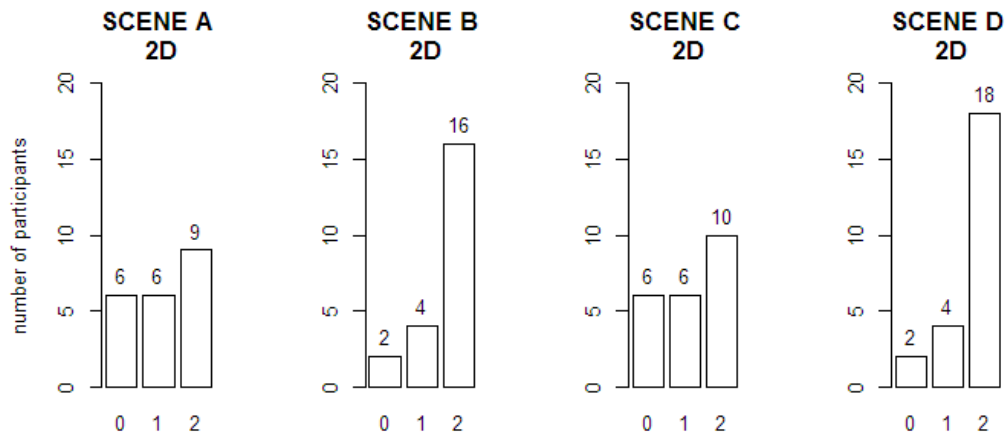


Figure 4-11. Boxplots showing frequency distribution of participant stereoacuity for Scenes A to D, for 2D (top) and 3D groups (bottom). Number of participants are plotted against Randot Test scores 6/10 to 10/10 (y axis).

4.5.2 A priori lidar knowledge and experience per 2D/3D group

Rank of lidar/laser-scanning description



Rank of lidar/laser-scanning experience

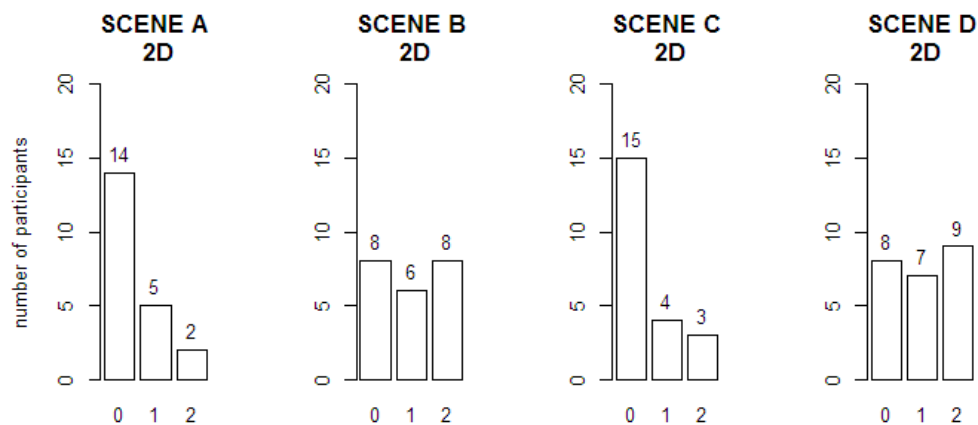
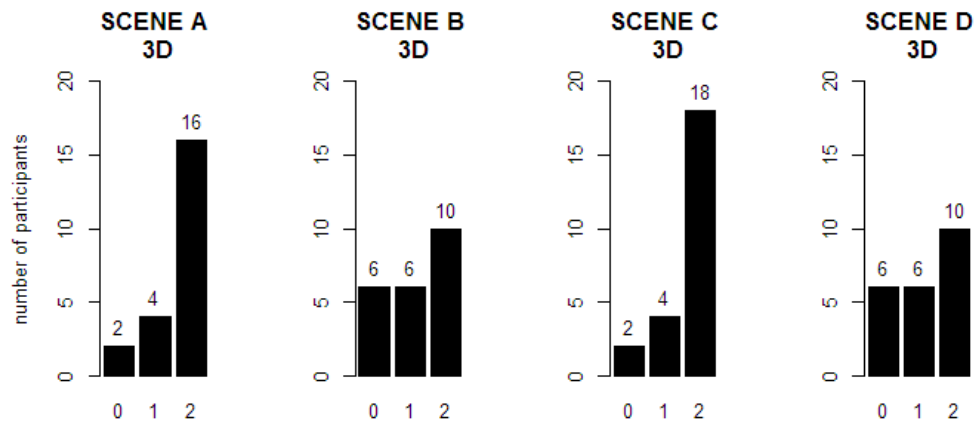


Figure 4-12. 2D participants' levels of lidar/laser-scanning knowledge (top) and experience (bottom). For knowledge, 0 = novice, 1 = informed, 2 = expert. For experience, 0 = novice, 1 = informed, 2 = expert.

The *a priori* expertise of the participants was measured (section 4.2.2) and Figure 4-12 and Figure 4-13 show the number of 2D and 3D participants with novice, informed, and expert knowledge (top rows) and experience (bottom rows), for each scene. A comparison of 2D and 3D groups (white bar charts in Figure 4-12 vs. equivalent black bar charts in Figure 4-13) shows differences in distributions for lidar knowledge and experience. For example, the Scene C 2D group is skewed towards novice experience, compared to the 3D group.

Rank of lidar/laser-scanning description



Rank of lidar/laser-scanning experience

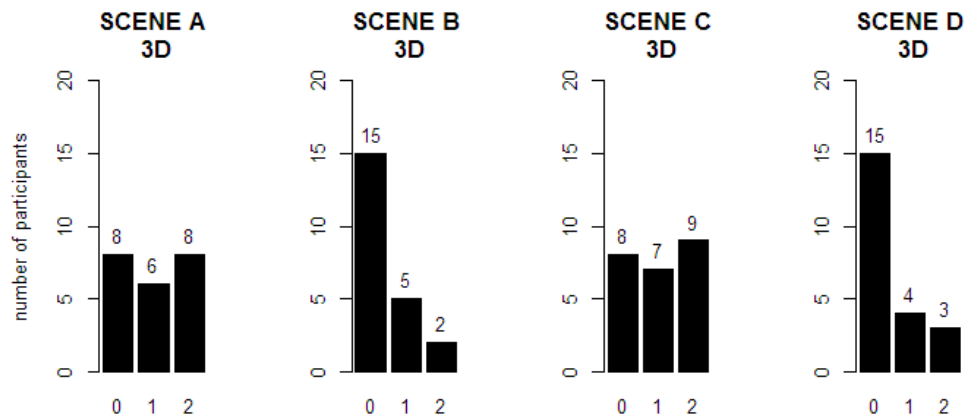


Figure 4-13. 3D participants' levels of lidar/laser-scanning knowledge (top) and experience (bottom). For knowledge, 0 = novice, 1 = informed, 2 = expert. For experience, 0 = novice, 1 = informed, 2 = expert.

4.5.3 Technology habits per 2D/3D group

In the written survey, all participants were questioned about how often they use different HCI devices while playing computer games. The results are shown in Figure 4-14 and most participants (80%) reported that they had previously used a gamepad to play computers games, with 16% using them between at least once a week to least once a day. Sixty-six percent of the participants had used a Nintendo Wii remote (visit www.nintendo.co.uk , accessed 10-09-15) and

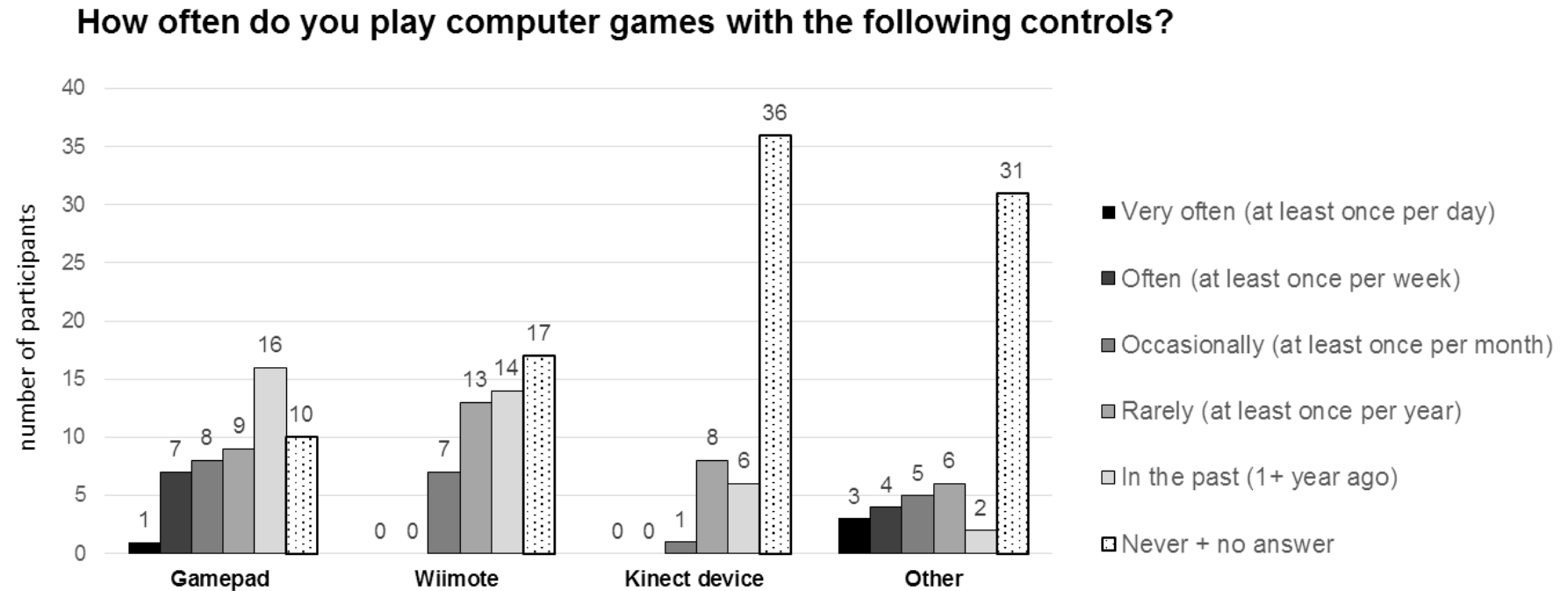


Figure 4-14. Summary of frequency of navigation device usage, for all participants.

29% had used a hands-free Microsoft Kinect device (Microsoft Corporation, 2015). Alternative technologies included a Leap Motion (www.leapmotion.com, accessed 09-09-15), stylus, joystick, touch screen, keyboard and mouse. A gamepad was used in the experiment, so the data are shown in Figure 4-15 (note changes in axes). Across each method, for each scene, there is variation in the frequency of gamepad use.

The participants were also asked about how often they viewed 3D displays, and how effective they found these to be. Results are summarised in Figure 4-16 for all volunteers. Figure 4-17 and Figure 4-18 display the frequency that 3DTV or 3D cinema displays were viewed by participants of the main experiment. Overall, the 2D and 3D groups rarely used this visualisation technique, but when they did they experienced strong 3D depth (ranked 4, where 1 is not at all effective, *i.e.* no extra depth in 3D; 5 is very effective, *i.e.* very strong 3D depth).

4.6 Summary

Chapter 4 presented an insight to participant background, including overall demographic groups and stereo test scores. Characteristics and technological habits of each trial group were also reported for each of the scenes.

Furthermore, since the participants were screened based on their stereo score, the relationship between gender and stereoacuity was analysed. This confirmed that the screening process did not cause gender bias in the sample of participants who took part in the full experiment.

How often do you play computer games with: a gamepad?

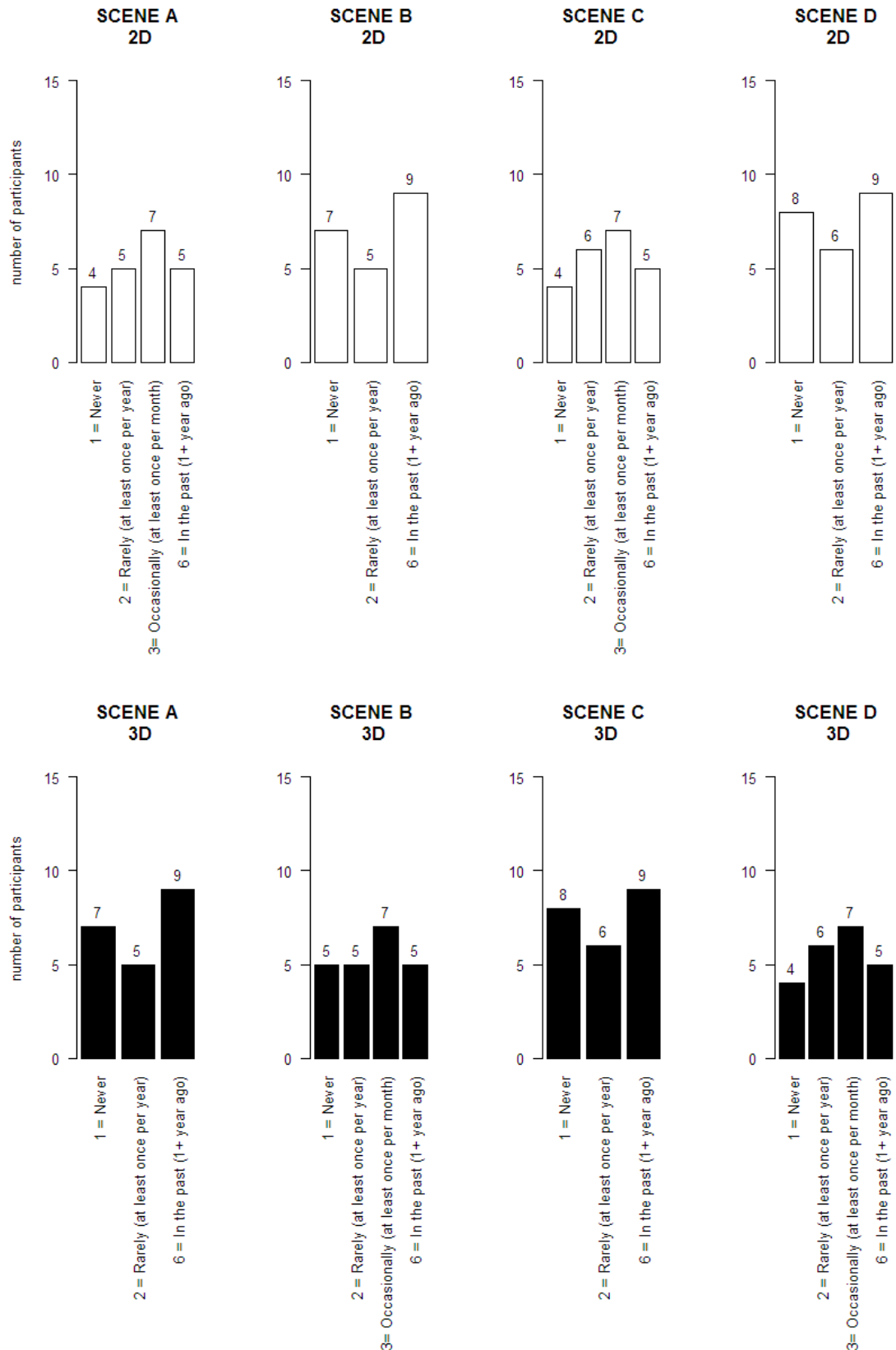


Figure 4-15. Frequency that 2D (top) and 3D (bottom) groups of participants, for each scene, use a gamepad device during computer gaming.

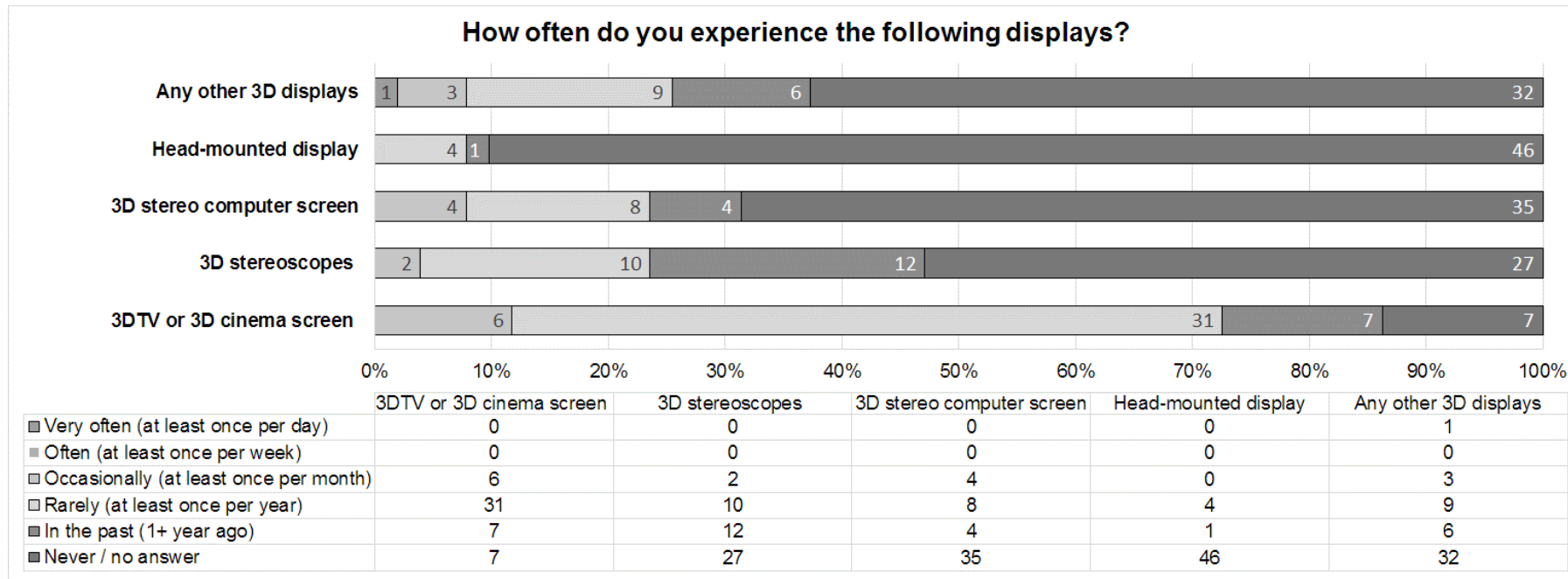
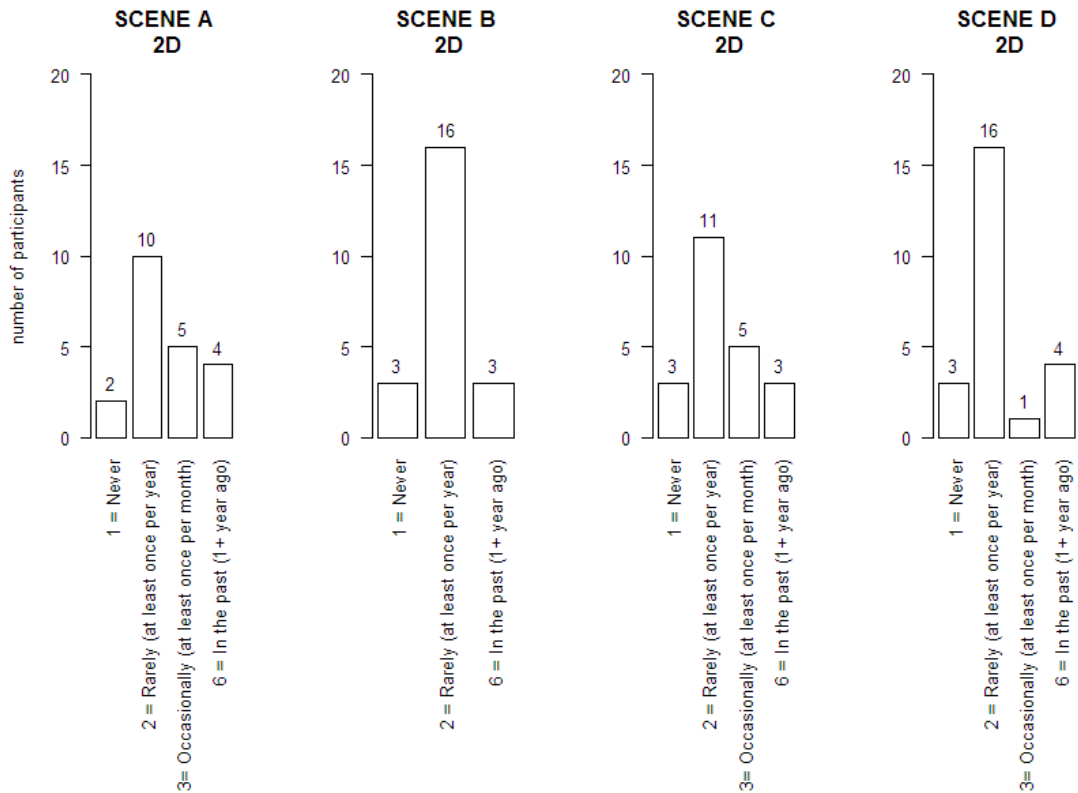


Figure 4-16. Summary of frequency of 3D display usage, for all participants.

How often do you experience: 3DTV or 3D cinema screen?



If you have ever experienced a 3DTV or 3D cinema screen, how effective* did you find them?

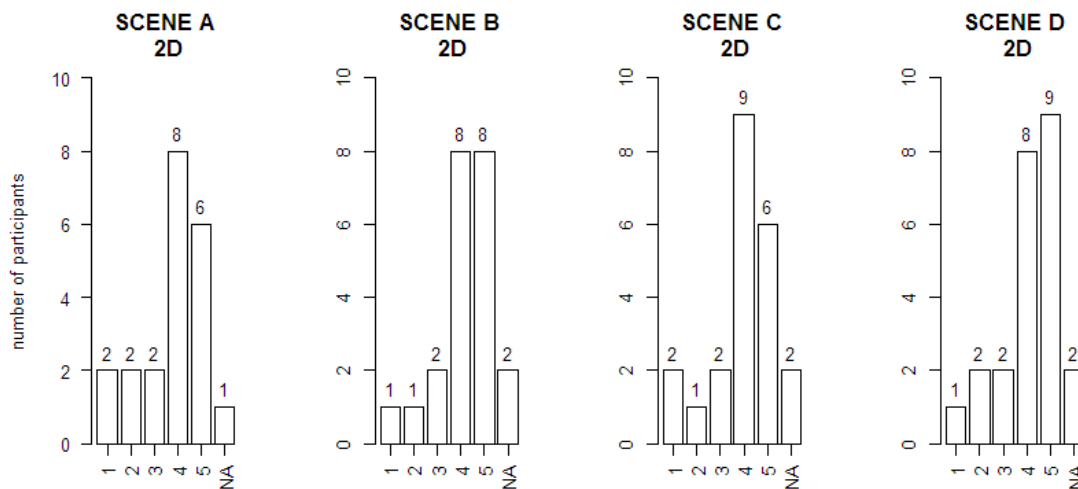
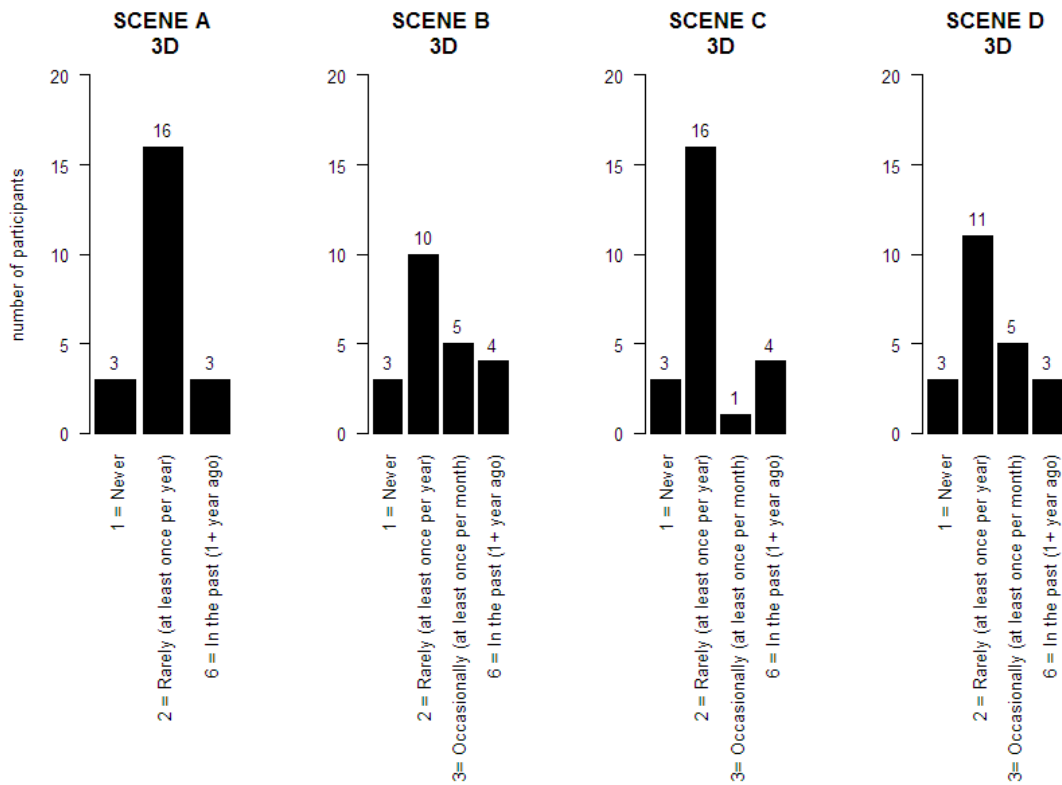


Figure 4-17. Frequency that 2D groups of participants, for each scene, experience 3DTV or 3D cinema screen (film, sporting event, etc.). *Effective, meaning you felt that the 3D experience gave added depth to the images. Response ranked from 1, Not at all effective (not extra depth in 3D), to 5, very effective (very strong 3D depth).

How often do you experience: 3DTV or 3D cinema screen?



If you have ever experienced a 3DTV or 3D cinema screen, how effective* did you find them?

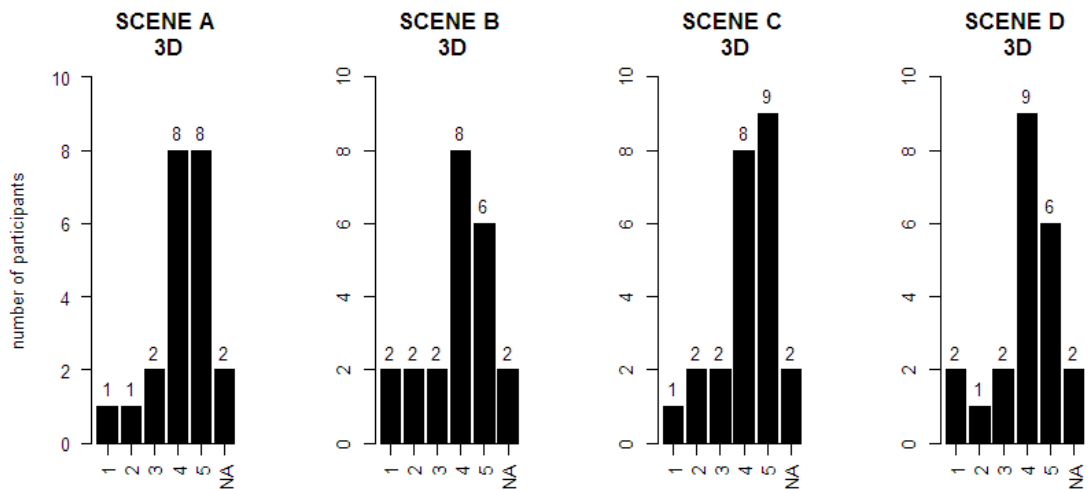


Figure 4-18. Frequency that 3D groups of participants, for each scene, experience 3DTV or 3D cinema screen (film, sporting event, etc.). *Effective, meaning you felt that the 3D experience gave added depth to the images. Response ranked from 1, Not at all effective (not extra depth in 3D), to 5, very effective (very strong 3D depth).

5.

Measurement

5. Measurement

5.1 Introduction

The measurement task was undertaken by participants to evaluate differences in 2D- and 3D-derived linear measurements from two different features. This chapter deals with the method used to develop of the task, its results, and finishes with a discussion of the findings in relation to other work. The aim of the task was to address Research Question 1, below.

RQ1 Is there a significant difference in linear measurement of lidar point cloud features derived from 2D and 3D visualisations?

RQ1.1 How precise are point cloud measurements made in 2D, in comparison to those made in 3D for (a) a planar feature and (b) a volumetric feature?

RQ1.2 Is there a significant difference between point cloud measurement made in 2D, in comparison to those made in 3D for (a) a planar feature and (b) a volumetric feature?

This was tested on two point clouds, Scenes A (planar feature) and Scene B (volumetric feature), which represent a house and a vegetation canopy, respectively. For each scene, the distribution of 2D and 3D results were compared to determine whether there was a significant difference between the methods (RQ1). Participant comments on perceived performance and preference were also extracted from the audio recordings to provide additional qualitative data. RQ3, restated below, reflects on the effectiveness of the novel method in answering RQ1.

RQ3 How effective is the methodology at comparing 2D vs. 3D visualisation of lidar point clouds? (**RQ3.1**) for the Measurement Task

The relevant research questions are sign-posted in the discussion section, to clearly link the observations of the experiment back to the crux of the research investigation.

5.2 Method

In addition to the general experiment method, the each task required further specific development, which is describe here, for the measurement task.

5.2.1 Task development

In order to widen the application of the study, the 2D and 3D manual tasks required visual stimuli (*i.e.* point clouds) of differing structure and dimensionality (Garg *et al.*, 2002; Jones *et al.*, 2008). The pilot study confirmed that two contrasting point cloud structures could be explored within a reasonable time, when taking into consideration the wider experimental set-up.

Site Selection

Two AOIs were identified from the Bristol dataset to represent planar and volumetric features at a geographical scale. These AOIs are referred to as Scene A (planar) and B (volumetric), and are summarised in Figure 5-1 and Figure 5-2 (also shown in greater detail Appendix C). The density of the ground points is sparser than the other points owing to the processing algorithm that was used, as previously explained in the method (Chapter 3).

Scene A (approx. 30m x 34m) contains a 2-storey house and its garden. The point cloud had been classified into ground, building, and vegetation (explained in section 3.2.1), the points representing a small tree were removed so that the participant was only focusing on points representing the building (and ground). The dataset included flightline overlap, which meant that, at Scene A, points were available from an ascending and descending flight path. This was deemed appropriate for the task as the aim was to measure the edge of the rooftop and to understand the angular shape of the point cloud (*versus* Scene B). The presence of two flightlines will have increased the point density of the roof.

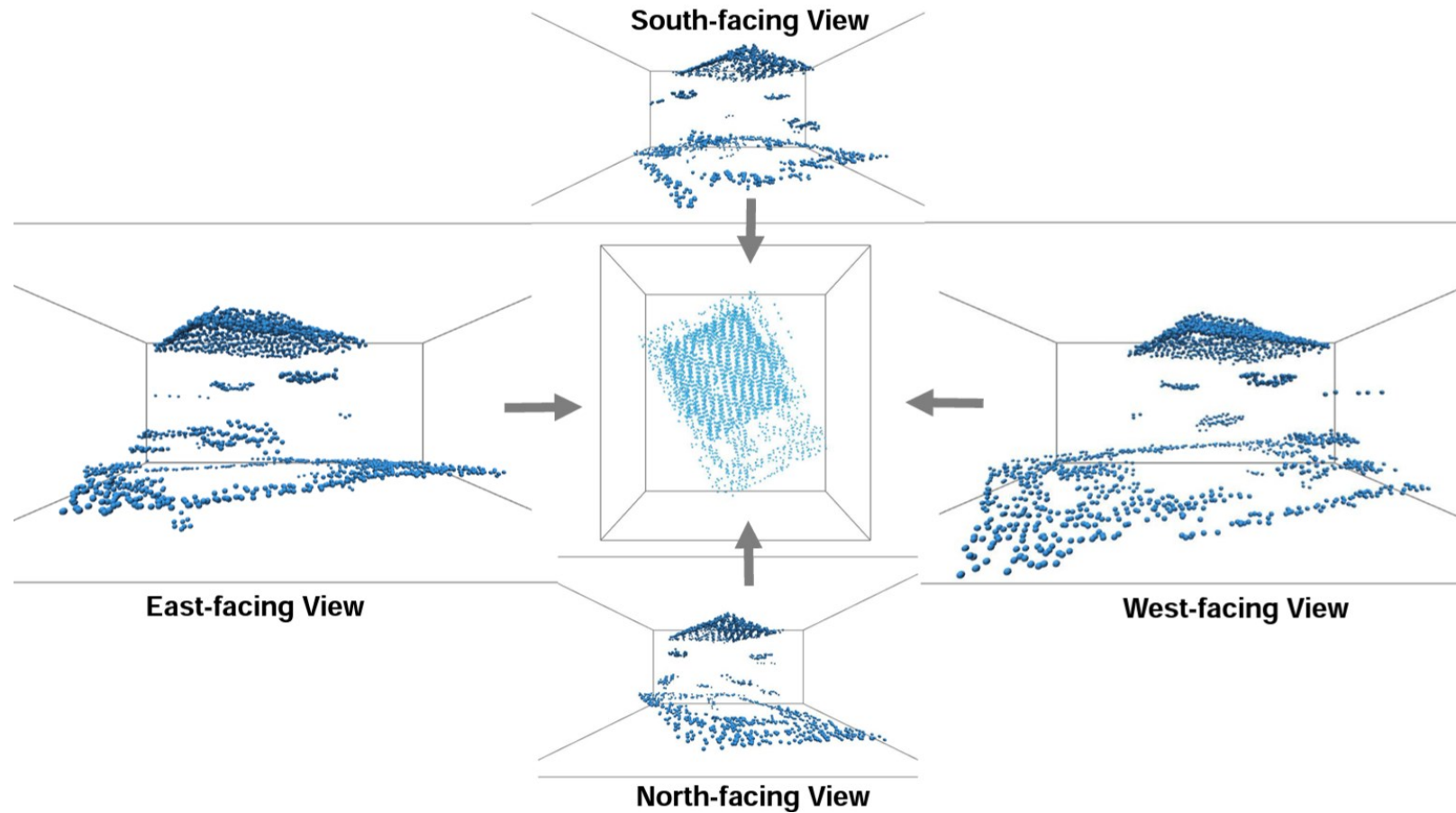
Scene A: planar building, for measurement task

Figure 5-1. Plan and side views of Scene A point cloud, which measures (width x length x height) 29.78m x 33.7m x 15.1m and is made up of 824 ground points and 1323 building points. Airborne lidar data © Airbus Defence and Space Ltd. (2013a)

Scene B : volumetric vegetation canopy, for measurement task

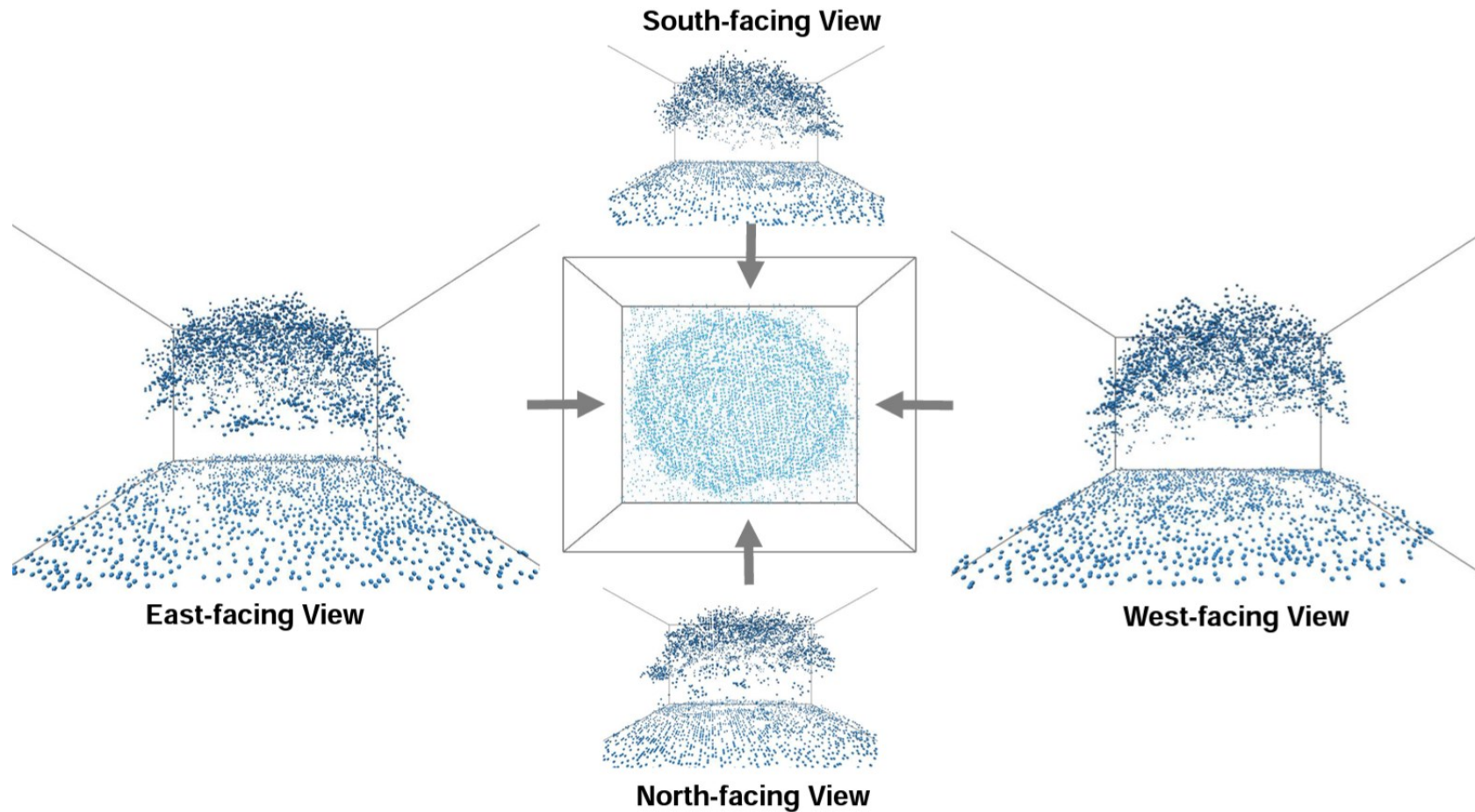


Figure 5-2. Plan and side views of Scene B, which measures (width x length x height) approx. 44m x 38m x 25m. The points cloud is made up of 1699 ground points and 2666 high vegetation points. Airborne lidar data © Airbus Defence and Space Ltd. (2013a).

Scene B (approx. 45m x 25m) consists of a group of trees, whose canopy is penetrated by the lidar, creating a volumetric arrangement of points. The location of the trees, at a roundabout (Appendix C), made it easy to isolate the vegetation from the buildings. Again, the points that did not belong to the feature of interest were removed to ensure the participants were solely focused on the feature. In Scene B, this meant that the vegetation and ground points were included in the visualisation, but building points were omitted.

Hypothesis

It was expected that a planar feature, such as a building, would be easy to measure in both 2D and 3D because it already has linear elements within the scene, which users could use as a reference edge. A volumetric structure, such as vegetation, would be more irregular and therefore more cognitively demanding to understand in 2D and 3D, compared to a planar feature. However, it is assumed that volumetric features may benefit from dynamic and stereoscopic projections, as suggested by Garg *et al.* (2002) and Seipel *et al.* (2012).

Measurement technique

The first task that the participants undertake is referred to as the measurement task, although it involves the user-generation of a vector whose the length is calculated post-experiment. The linear measurement task was chosen because it relies on the manual selection of two points in order to fit a vector, from which distance can be calculated. The task requires the participants to determine different breakpoints in Scene A and Scene B. No measurement calculations are carried out by the participants, only selection of points from the on-screen data to create a vector. No tools or aides were available to assist participants during the task; only the point cloud was shown and a crosshair could be toggled on/off by the participants to help them target points.

Measurement of feature height was discounted as a potential estimate for the participants because ALS data can lead to under- or over-estimation of feature

height, as shown in Figure 5-3. Zimble *et al.* (2003) also highlight this drawback in a forestry context, whereby the tallest part of the tree can be missed by the lidar points, depending on the location of the intermittent hits. Participants would be disadvantaged by the vertical occlusions as a result of interceptions during acquisition (*i.e.* building roof prevents ground hits). A horizontal linear measurement was deemed more appropriate for the test, as breaklines can be determined subjectively, although this too carries inherent uncertainty because of the intermittent lidar returns. Point cloud density would also have an influence on the outcome.

Owing to the acquisition time of the ALS data (in 2008), no reference measurements were used in the study, against which the accuracy of each visualisation method could be compared. With ALS surveys, a ground survey must be timed to coincide with the data acquisition.

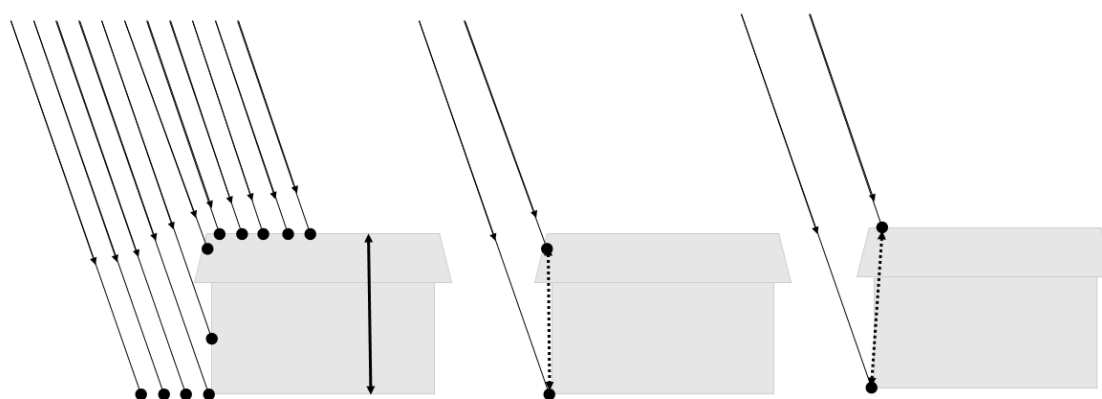


Figure 5-3. Issues with ALS height measurement, directly from point cloud. Left - Lidar points available for measuring height of building (solid vertical arrow), middle – measuring from point above ground point gives an underestimated height, right – measuring from the correct height, at a different xy position leads to overestimation of the feature height. Author’s own image.

The measurement tools were designed around the gamepad, in addition to its navigation functionality. During the task, an output file was generated for each participant, in which the selected point locations were recorded alongside the time-of-button-press. The xyz coordinates of the participants’ first (P_1) and second (P_2) point selections were recorded during the trial and the measurements were calculated outside the visualisation system. The raw data

text file was imported into Excel (Microsoft Corporation, 2012a), where the Pythagorean Theorem was applied (Appendix K) to calculate distance in a three-dimensional coordinate system. Equation 2 was used to calculate the three-dimensional distance between selected points

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Equation 2

Where:

$d(P_1, P_2)$ = distance between Point 1 and Point 2,

Values x_1, y_1, z_1 = xyz coordinates for Point 1

Values x_2, y_2, z_2 = xyz coordinates for Point 2

5.2.2 Experiment instructions

The participants were presented with their first scene (Figure 5-4, either Scene A or B) in either 2D or 3D, depending on the random order issued to each individual. Participants were shown the different aspects of their first scene (plan view, north-facing, etc., as shown in Appendix C)

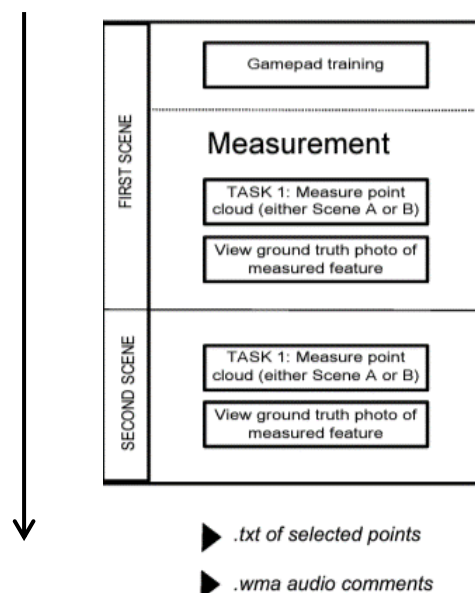


Figure 5-4. Elements of the measurement task in relation to first and second scenes. Vertical arrow indicates passage of time. Black triangles denote file outputs from the task.

Gamepad

At the beginning of the measurement task, participants were told that they would measure the features within the scenes by selecting points, using the gamepad device. After having undertaken gamepad training during stage two of the experiment (section 3.4.3), gamepad instructions relating to the measurement task (Figure 5-5) were shown and explained to the participant (Appendix H).

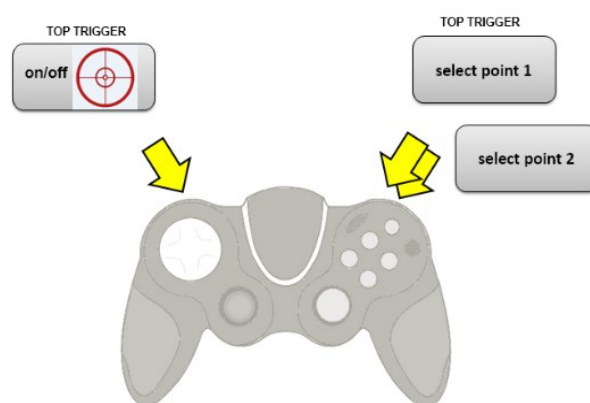


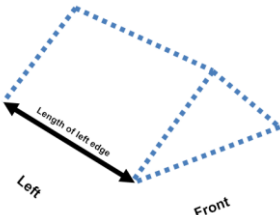
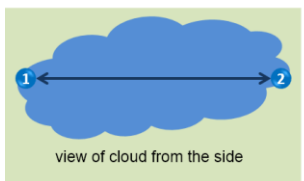
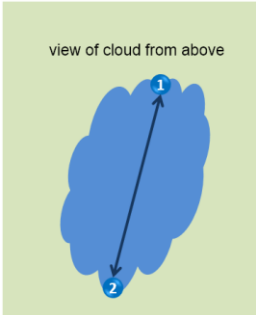
Figure 5-5. Point Selection Training – guidelines shown to participants, indicating gamepad buttons used to toggle on/off crosshair target icon and select 1st and 2nd points.

Measurement

Measurement instructions were read aloud by the researcher (Appendix H) and accompanying schematic diagrams were shown to ensure participants were measuring the same element of each feature (summarised in Table 5-1). Ultimately, the decision of point selection was carried out by each individual. Once the participants had been briefed about the task, they were allowed to identify and select their two points of choice.

Figure 5-6 illustrates the set-up of the VR theatre during this task and the steps taken by each participant for Scene A and B. During navigation, participants were told to keep the on-screen crosshair turned off. Once two points were selected, a line was generated and the task ended.

Table 5-1. Instructions told to participants during the measurement task. Verbal and visual directions were given.

| Instructions | |
|---------------------|---|
| General | <ul style="list-style-type: none"> • I'm going ask you now to take a measurement from this point cloud • I will show you where I'd like you to take the measurement from. • Look around and decide which two points you think they are. • The images are examples only; you can take the measurement from any angle. |
| Scene A | <div style="text-align: center;">  </div> <ul style="list-style-type: none"> • Measure the left edge of the feature. |
| Scene B | <p>Measuring the longest diameter (or longest spread)</p> <div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;">  <p>view of cloud from the side</p> </div> <div style="text-align: center;">  <p>view of cloud from above</p> </div> </div> <ul style="list-style-type: none"> • Measure along the longest diameter. • Measure the width of the feature at its fullest part • Select the two points that you consider to be the widest transect that cuts through that feature. |

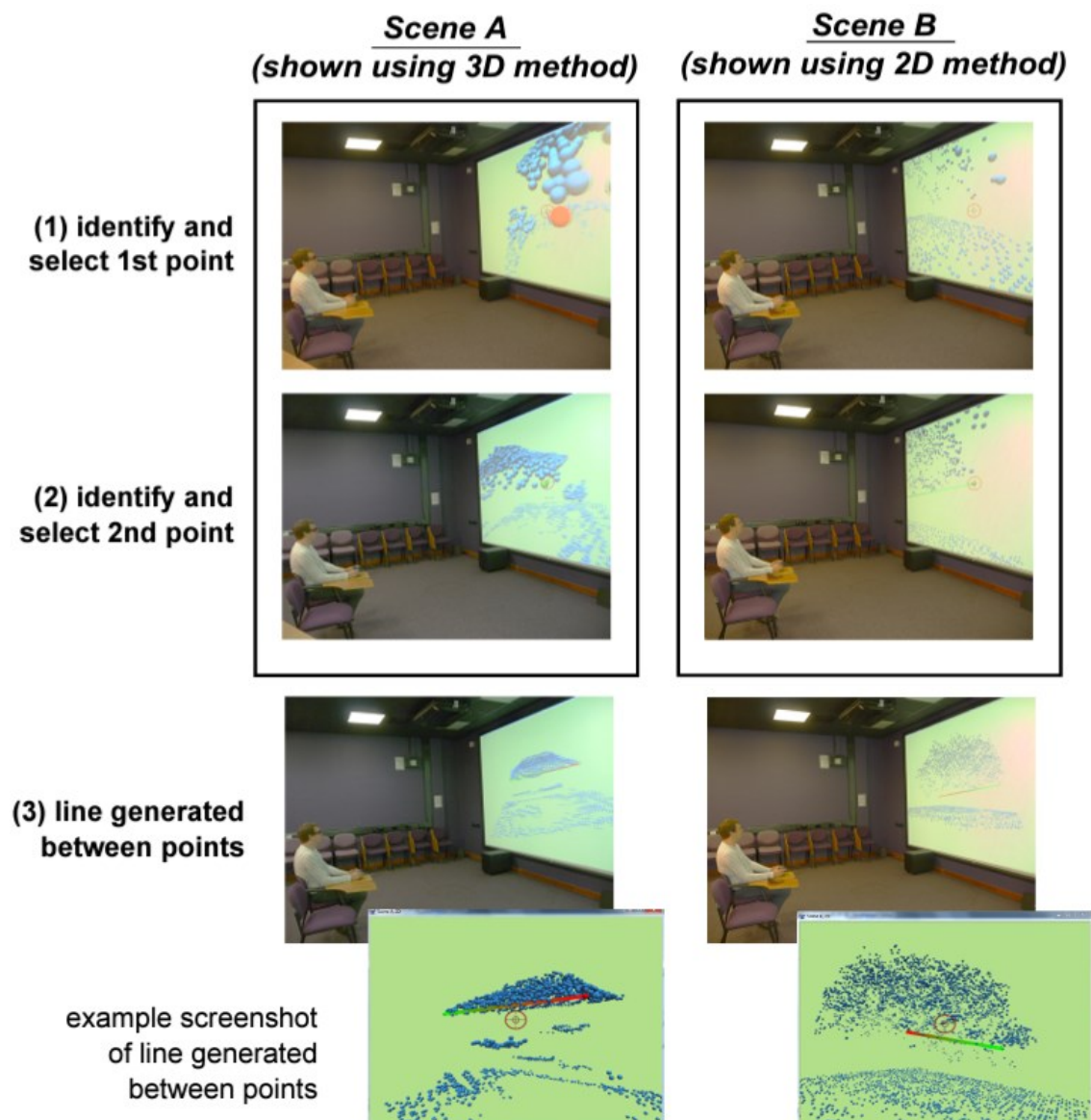


Figure 5-6. Photographs of experiment set-up during measurement task. Boxed images represent action carried out by participant. Other images show the resulting vector that is generated between the participants' two points – in situ and screenshot examples. Participant posed by model. Airborne lidar data © Airbus Defence and Space Ltd. (2013a).

Reference data

After completing a measurement, participants were shown aerial and street-level photography of the AOIs (shown in Appendix C) so that they could relate the point cloud representation to the real-world feature. This contextualisation step was a prerequisite for the following interpretation task.

The experiments were recorded with a dictaphone throughout, providing qualitative results to support or counteract (Andre and Wickens, 1995) the quantitative results.

5.2.3 Data Analysis

From the experiment, measurement results were available from participants who had measured Scene A in 2D and those who had measured it in 3D. The same was true for Scene B and the following statistical tests were carried out on these data. Participant feedback was also analysed for the measurement scenes to provide a narrative to the 2D and 3D results.

Descriptive Statistics

For Scene A, the participants who were exposed to the 2D display were different to those who experienced the 3D stereo display. This was also the case for Scene B. Therefore, two sample populations existed for each Scene, which were made up of different participants (2D group \neq 3D group). The unpaired approach was vital to eliminate any learning effect that would have resulted from the same participants carrying out both methods. For example, if a participant were to carry out measurements in 2D on one of the scenes, followed by measurement of the same scene in 3D, the second measurement may have been biased by the knowledge and experience gained during the first measurement. For the measurement task, any comparison between the 2D and 3D methods are considered *unpaired* because results are compared between different participants. If data distribution fits a normal curve (as described in section 3.5.2), parametric tests are carried out. Confidence interval tests relate to the assumed normal frequency distribution, illustrated by the histogram in Figure 3-20, in Chapter 3. Box plot graphics are also used to further investigate the distributions of the measurements generated by the 2D and 3D groups of participants. Figure 3-21 shows the elements of a boxplot, in relation to a histogram. Figure 5-7 illustrates potential precision of the observed measurement values (and accuracy scenarios in relation to reference values). In this study, owing to the absence of ground reference data that coincides with ALS survey, the accuracy is not formally assessed. If ground survey measurements were available, this could also be analysed, alongside precision.

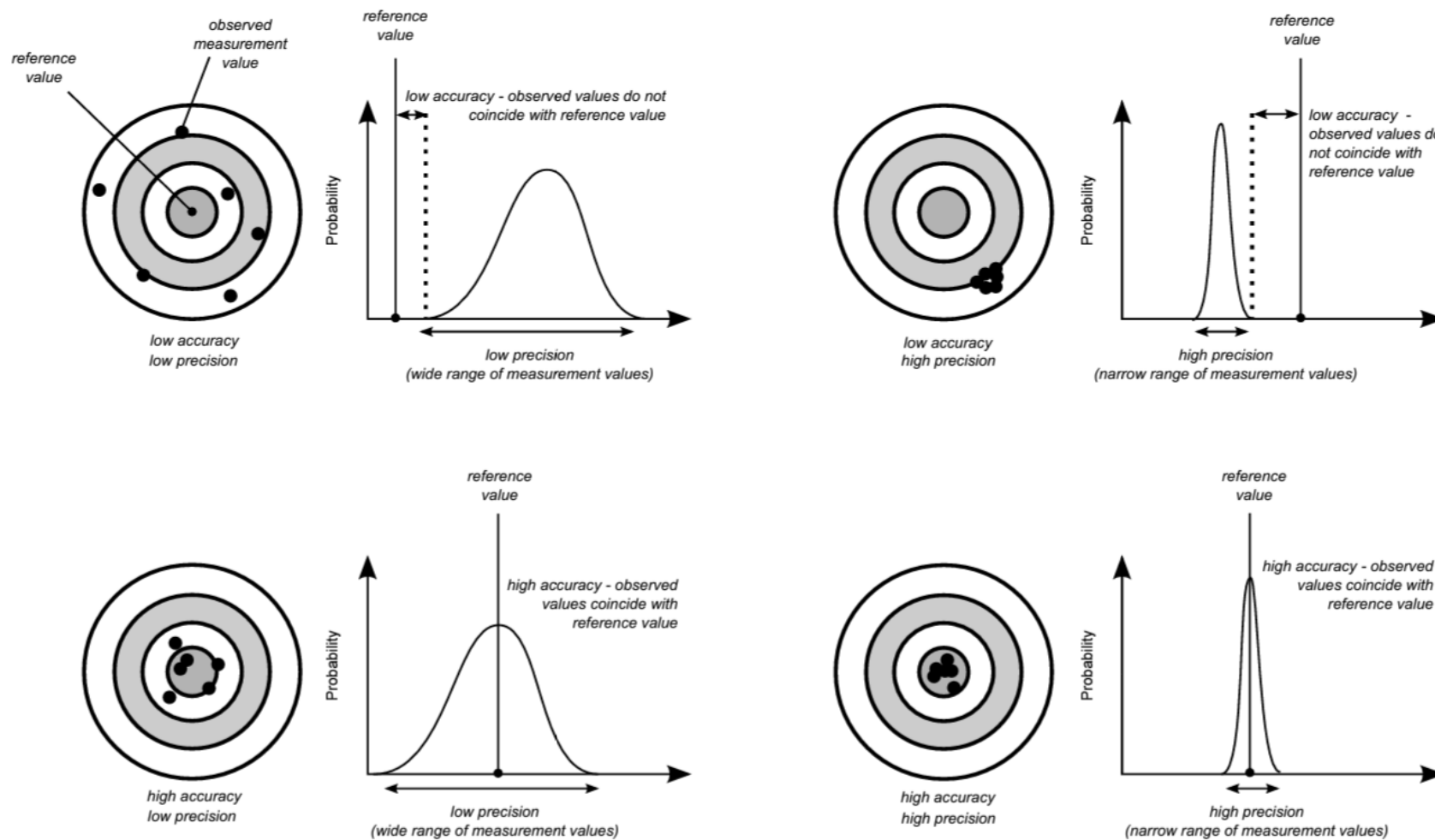


Figure 5-7. Schemas comparing precision and accuracy of observed values using a dart board and the approximate equivalent histogram shapes (where x axis is the observed measurement values and y axis is the frequency distribution of the sample group). ALS data were acquired in 2008, so time of flight ground survey was not available for an appropriate reference value. Author's own diagram.

Prior to comparing the 2D and 3D results, firstly, a quantile-quantile plot (or Q-Q plot) is created in R using the `qqplot` function (R Core Team, 2014). This illustrates how the distribution of observed sample measurements align with the theoretical normal distribution of measurements (Gilbert, 1987). Figure 5-8 illustrates an example of a Q-Q plot; the solid line represents the expected normal results and the dots are the observed observations. If the dots are aligned with the line, this indicates that the observed measurements exhibit a normal distribution.

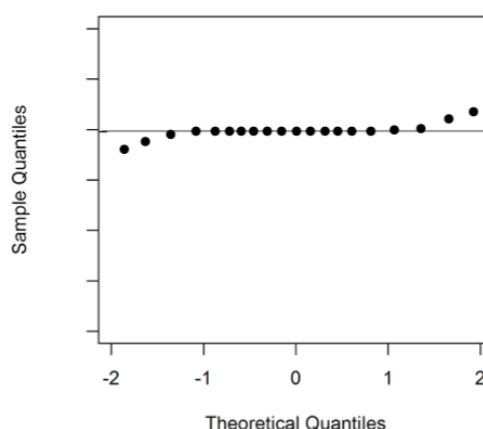


Figure 5-8. Example of a Q-Q plot showing participant measurements (dots) relation to normal distribution. Y axis = measurement values (m); if all participants measure the same distance, the dots would lie in line with each other.

In addition to this graphical Q-Q plot check for agreement, the Shapiro-Wilk test (Shapiro and Wilk, 1965) is carried out using the `shapiro.test(x)` R function (R Core Team, 2014), with updates from Royston (1995), to determine the departure from normality of each sample. The Shapiro-Wilk test is suitable for smaller sample sizes (less than 50) and tests the observed data against the theoretical normal distribution (based on same degrees of freedom and standard deviation values). The test returns values between 0 and 1, where 1 indicates agreement between the two datasets. If the sample does not meet assumptions of normality, meeting the H_1 hypothesis, below, non-parametric tests must be employed during further statistical tests (Gilbert, 1987).

- H_0 – There is no significant difference between actual and theoretical samples; the population has a normal distribution.
- H_1 – There is a difference between actual and theoretical samples; the population does not have a normal distribution.

Comparison of 2D vs. 3D measurement values

The similarities between the observed 2D versus observed 3D results are explored by plotting the results against one and other in a Q-Q graph. A perfect correlation would show a 1:1 like-for-like relationship, as shown in Figure 5-9.

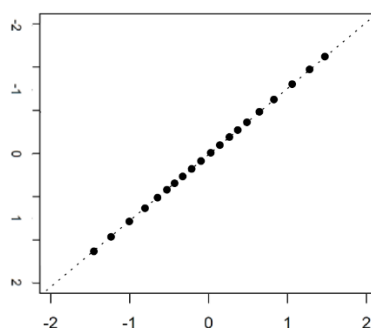


Figure 5-9. Example of a QQ plot where 2D (x) vs. 3D (y) measurement results are perfectly correlated.

This is explored further by carrying out a Mann-Whitney Test (Mann & Whitney, 1947) in R, otherwise known as a Wilcoxon signed rank test (with continuity correction). The unpaired test is appropriate for the comparison of two independent groups that do not exhibit normal distributions.

- H_0 – There is no significant difference between the mean measurements of the 2D and 3D groups.
- H_1 – There is a difference between the mean measurements of the 2D and 3D groups.

Quantitative and qualitative audio analysis

The audio file transcriptions were searched, using Nvivo software (QSR International, 2013), for qualitative comments relating to the Measurement Task. Themes were also identified within the transcriptions by allocating selected text to user-defined nodes or categories. Further comments from the Feedback section of the experiment were also extracted, relating to the participants' perception of measurement accuracy based on different conditions.

5.3 Results

The measurement task results contribute to answering research question 1.1 by describing the point cloud feature measurements for a planar and volumetric feature.

5.3.1 Planar feature (Scene A) – roof edge length

Twenty-one participants carried out the planar measurement in 2D and the 3D group was made up of 22 participants. Three sets of measurement datasets did not record to the output files, so the results of two 2D participants and one 3D participant were discounted.

Frequency distributions for 2D & 3D planar measurements

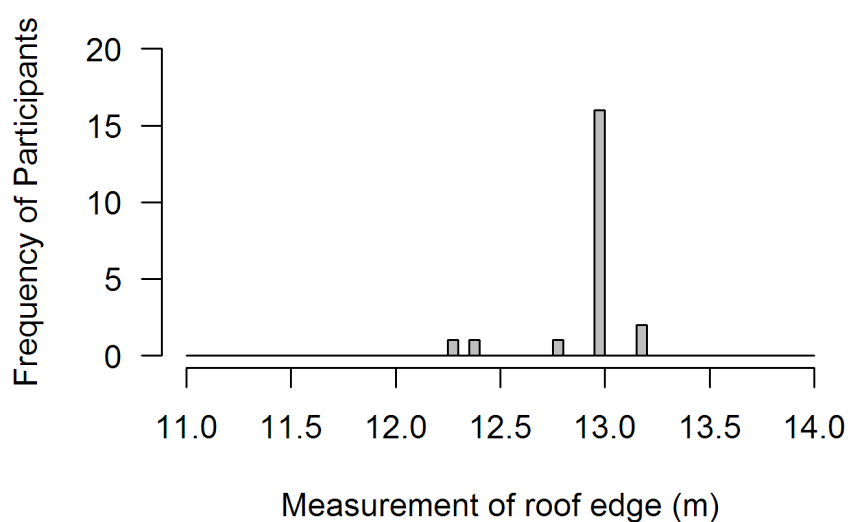
The planar measurement results are summarised in Table 5-2 and show that the 2D and 3D group both have the same median (12.98m), but different means (12.92m and 12.88m, respectively). The histograms Figure 5-10 show that the data are negatively skewed, which suggests the distribution is not a normal Gaussian shape.

Table 5-2. Distribution of measurements in the 2D and 3D groups for Scene A's roof edge.

| Distribution | 2D group (m) | 3D group (m) |
|--------------|--------------|--------------|
| Minimum | 12.27 | 11.33 |
| 1st Quartile | 12.98 | 12.98 |
| Median | 12.98 | 12.98 |
| Mean | 12.92 | 12.88 |
| 3rd Quartile | 12.98 | 12.98 |
| Maximum | 13.18 | 13.17 |

The measurement range for 2D (0.91m) is approximately half that of 3D (1.84m), which suggests that 2D technique derives more precise, although not necessarily accurate, results (Figure 5-7). However, the boxplots in Figure 5-11 show clearly that the larger range in measurement values in the 3D technique is caused by outliers, of which there are relatively few.

Scene A: 2D group (n=21)



Scene A: 3D group (n=22)

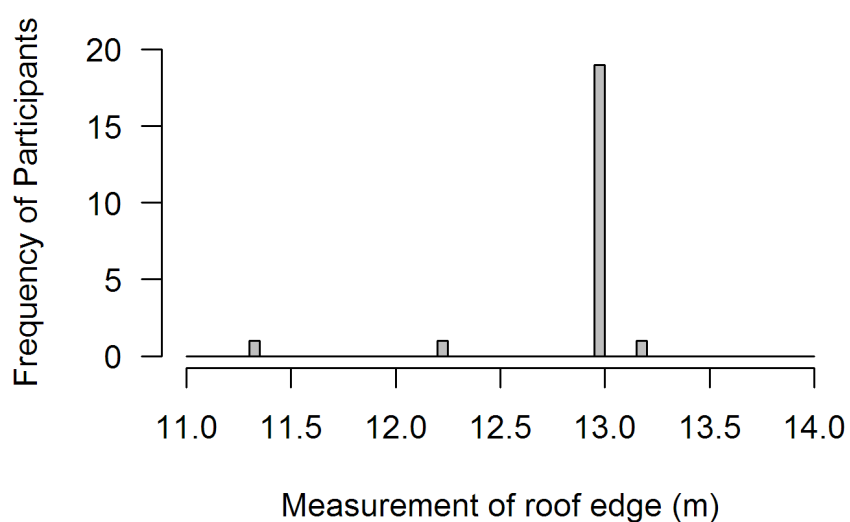


Figure 5-10. Histogram of Scene A roof edge measurements for 2D participants (top, n = 21) and 3D participants (bottom, n = 22). Bin size = 0.1m.

In Figure 3-21, IQR is shown to represent the result range of 50% of the participants about the mean. The boxplots in Figure 5-11 have an absence of this box element, which indicates that over 50% of both the 2D and the 3D group had a median value of 12.98m. When measuring a planar feature length, fewer of the participants from the 3D group deviated from the median value, compared to the 2D group. However, the 2D group had a smaller range of measurements, compared to the 3D group. These results show that the 2D method generates a more precise result. This is observed more clearly in the Q-Q plots in Figure 5-12 (bearing in mind 2D: $n = 21$, 3D: $n = 22$), which display the observed quantiles against theoretical for each method. In order to determine whether parametric or non-parametric techniques should be used when directly comparing the 2D and 3D data, the normality of each of their distributions were investigated. The black dots are the observed measurements that were carried out by the participants and the solid line represents the theoretical distribution for that sample of the population.

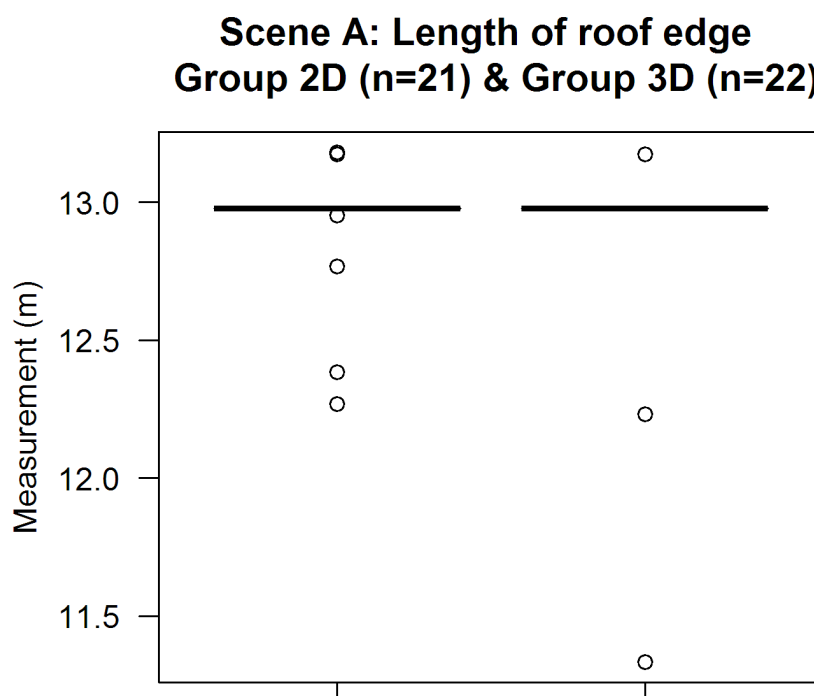


Figure 5-11. Boxplots of Scene A measurements for 2D participants (left, $n = 21$, y range from 12.27m to 13.18m) and 3D participants (right, $n = 22$, y range from 11.33m to 13.17m).

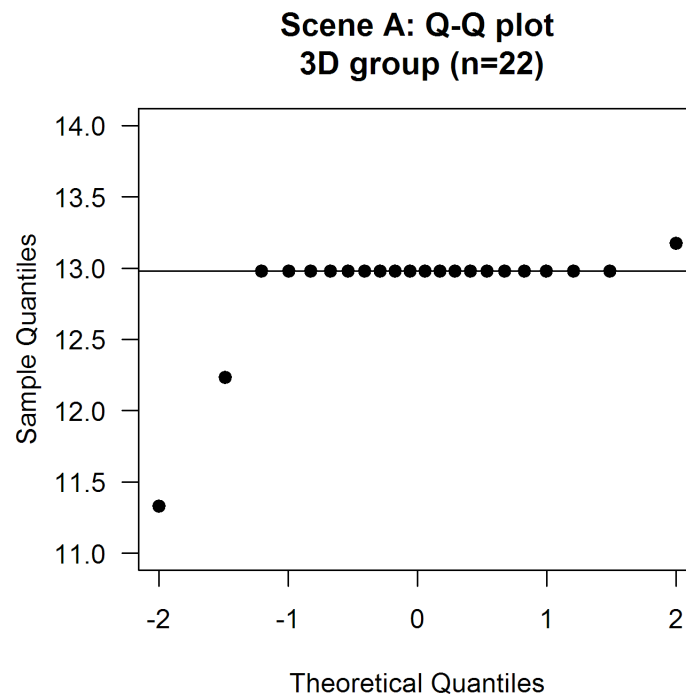
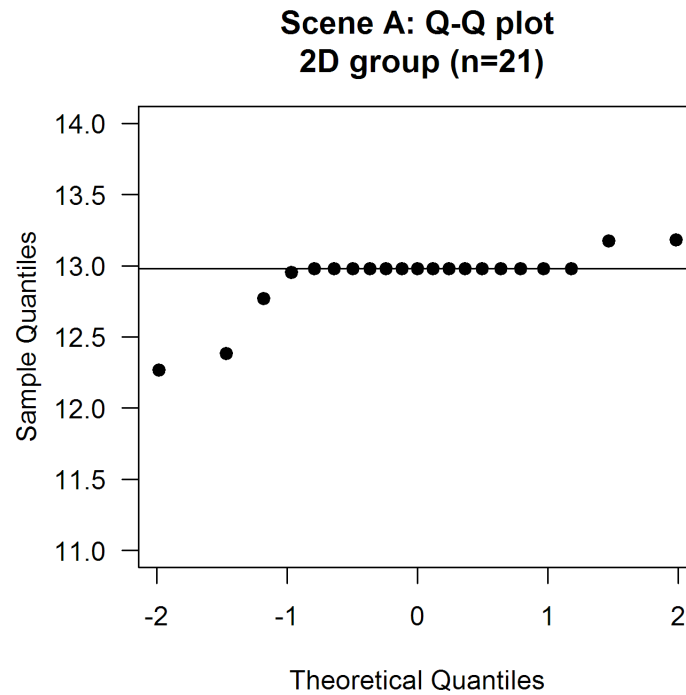


Figure 5-12. Theoretical vs. Sample quantile plots for Scene A 2D (top) and 3D (bottom). X axis: 0 = mean (50% frequency) and deviations from this represent one quartile (or 25%). Y axis = measurement results (in m) made by participants. Each point represents an observed measurement made by one participant. Solid line denotes the theoretical normal distribution.

When referring to Figure 5-12, the stereoscopic 3D measurement of the roof edge, compared to a standard flat 2D display, had fewer outliers around the expected distribution. The 2D plot (top) indicates that 6 participants (28.6% of total) are unaligned with the expected normal distribution (solid line), whereas the stereo group, Figure 5-12 (bottom), has 3 participants (13.6% of total) that do not fit the normal distribution line. Of the 2D group, 71% of participants achieved the correct length of 12.98m, compared to 86% of the 3D group. However, the measurement range is larger for the stereo method, suggesting the 2D technique generates more precise measurements from a planar feature.

Prior to comparing the probability distributions between the 2D and 3D methods, a test for normality was carried out on each set of results. Table 5-3 shows the Shapiro-Wilk Test results and both the 2D and 3D groups for Scene A have a p-value below 0.05. This means that neither groups match the normal distribution and non-parametric statistical tests must be carried out.

Table 5-3. Shapiro-Wilk normality test for 2D and 3D group, Scene A.

| Method | W value | P-value |
|--------|---------|----------|
| 2D | 0.5916 | 1.61E-06 |
| 3D | 0.3804 | 1.18E-08 |

H₀ – No difference between actual and theoretical (p.value ≥ 0.05)

H₁ - Difference between actual and theoretical. (p.value < 0.05)

Comparison of 2D vs. 3D planar feature measurements

The quantiles of the 2D and 3D groups are plotted in Figure 5-13, showing the agreement between results of the two methods. Although the plot shows samples deviating from the line of agreement, tied values are not obvious. As expected from Figure 5-12 findings, multiple points are overlapping at 12.98m (and elsewhere) from both groups.

To statistically test the difference between the two non-normally distributed groups, the Mann-Whitney test (Mann & Whitney, 1947; R Core Team, 2014) can be used to rank of scores of 2 independent variables. In R, this is carried out using an unpaired two-sample Wilcoxon signed rank test with continuity correction, using the following hypotheses:

H1.2a₀ – No difference between 2D and 3D planar measurements (p.value ≥ 0.05)

H1.2a₁ - Difference between 2D and 3D planar measurements (p.value < 0.05)

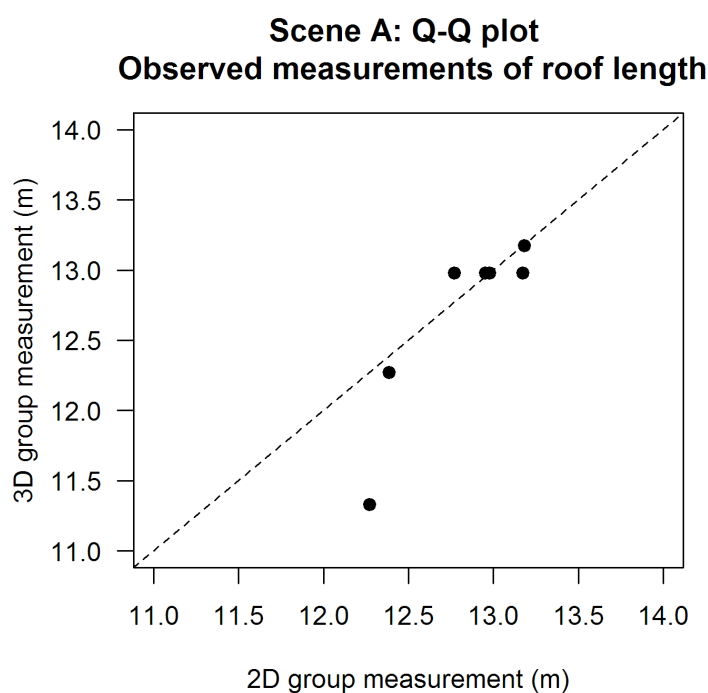


Figure 5-13. 2D vs. 3D quantile plots for Scene A, measurement of building roof edge. Dashed line represents 1.0 agreement, where 2D results= 3D results.

The estimated Mann-Whitney result ($U = 188.5$, $p\text{-value} = 0.2126$) rejects the null hypothesis that the means of the results are the same. This suggests that the rank totals for the 2D measurements are similar to the 3D ranks. In summary, under the experiment parameters, the 2D and 3D methods for measurement of a planar point cloud feature are not significantly different.

Participant perception of 2D & 3D measurement

The accuracy perception results for Scene A, regardless of the technique used, are shown in Figure 5-14. The positive skewness of this plot shows that participants generally felt that they measured Scene A accurately. The comments of the 2D and 3D groups, in Table 5-4 and Table 5-5, generally support this and reveal that the straight line and flat edge of the house was considered to be helpful when carrying out the measurement.

Q8: How accurately do you feel you measured Scene A?

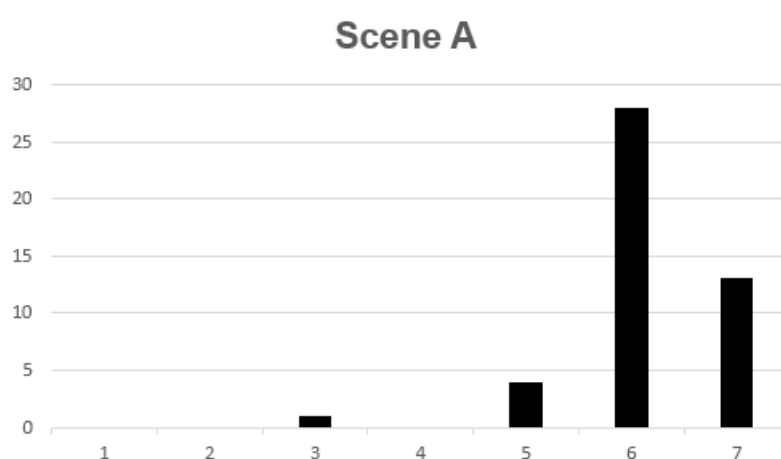


Figure 5-14. *General participant responses to Question 8 – “How accurately do you feel you measured Scene A?”. Scores were ranked on a scale of 1-7, where 1 = Not at all accurately, 7 = Very accurately. N.B. Values of 0 [no data inputted] not included on scale. Graph shows combined answers from 2D and 3D groups, n = 46 (inclusive of two participants whose quantitative results were discounted)*

Comments from both 2D and 3D participants regarding the difficulty of Scene A are found in the Appendix K. These comments of reflection were made during the feedback section of the experiment, normally approximately an hour after the measurement task was carried out. Generally, participants found that there were clear boundaries to the Scene A feature and the measurement was easier because it was along a flat edge. Participants who measured Scene A in 3D seemed particularly enthused about the straight edges, ‘*The sides were very apparent*’, ‘*with a definite start and end point*’.

Table 5-4. Further comments on Feedback Question 8, from participants who measured Scene A in 2D. Any comparisons are in reference to Scene B in 3D.

| P. no. | 'Further comment' on Q8 – Measurement of Scene A (and participant used 2D method) | Order |
|---------------|---|-----------------|
| P24 | I'd say I measured quite accurately in scene A, so probably a 6. | 1 st |
| P17 | That's quite a... that's just the roof of a building, all the points are clustered together. It might be [easier than Scene B in 3D] because of the actual thing that you're trying measure maybe, rather than the actual technique itself. | 1 st |
| P25 | I'd say that one's alright. I'd say... a 6 on that one. Still think there's room for improvement. I think... it's easier to choose specific points on that one, because you're picking a flat edge, so you choose the edge. And you kind of choose the point that comes to the forefront of that edge... and you just go with whatever you sort of think on that. | 2 nd |
| P26 | I think that was slightly easier, I'd say 6 again actually. | 2 nd |
| P50 | A – yeah, it was okay, wasn't it? Let's say 6. | 2 nd |

Table 5-5. Further comments on Feedback Question 8, from participants who measured Scene A in 3D. Any comparisons are in reference to Scene B in 2D.

| P. no. | 'Further comment' on Q8 – Measurement of Scene A (and participant used 3D method) | Order |
|---------------|---|-----------------|
| P34 | For Scene A I think I was very happy for my measurement for that. I wouldn't say 100%. It was helpful, the sort of STRAIGHT EDGE measurement. | 1 st |
| P43 | Um, scene A I think was pretty much spot on because it was a nice straight line | 2 nd |
| P12 | [Scene A] Oh yeah I was pretty confident with that. 6 I guess, because there's a chance I picked the furthest ones out that they actually weren't part of the roof. | 2 nd |
| P33 | A, I think I got fine, so go for a 7 on A. | 2 nd |
| P36 | I like the measuring in 3D, not the 2D. | 2 nd |
| P35 | [Scene A] For me? That's the best I could do, so, very accurately <laughter>. | 1 st |

Further comments are found in the Appendix K, which were made by participants during or before measurement of the roof edge. Members of the 2D and 3D groups expressed the intention to face the roof edge, so that they were perpendicular to measurement. One participant (P41) explained that although there were no axes in the GUI to denote orientation of the scene, the geometry of the building provided this information. Two participants (P28 & P29), both with high stereoacuity (Randot scores of 10/10, or 20 secs of arc), explicitly described their intention to exploit depth perception to identify the closest (corner) point to them. This demonstrates that they were able to take advantage of the extra dimension that was presented by the 3D method.

5.3.2 Volumetric feature (Scene B) - canopy diameter estimates

All participants estimated the longest diameter of the Scene B vegetation canopy, with 22 participants in the 2D group and 22 in the 3D group.

Frequency distributions for 2D & 3D volumetric measurements

The histograms in Figure 5-15 show us that the distribution of measurements made by both the 2D and the 3D group are positively skewed and do not follow a normal Gaussian, bell-shaped curve. The measurement range for the 2D group is 26.68m and 9.62m for the 3D group. The variation in results (from 8.85m to 35.53m) between the 2D participants suggests that the 2D technique is relatively imprecise (Table 5-6, p.126). The representation of the data as boxplots in Figure 5-16 show clearly that this is caused by outliers.

Nevertheless, the interquartile range of the 3D measurements is smaller. The thick black lines within the boxplots of Figure 5-16 show that the 3D median (34.44m) is higher than that of the 2D group (34.17m), suggesting that 2D participants underestimated the longest canopy diameter.

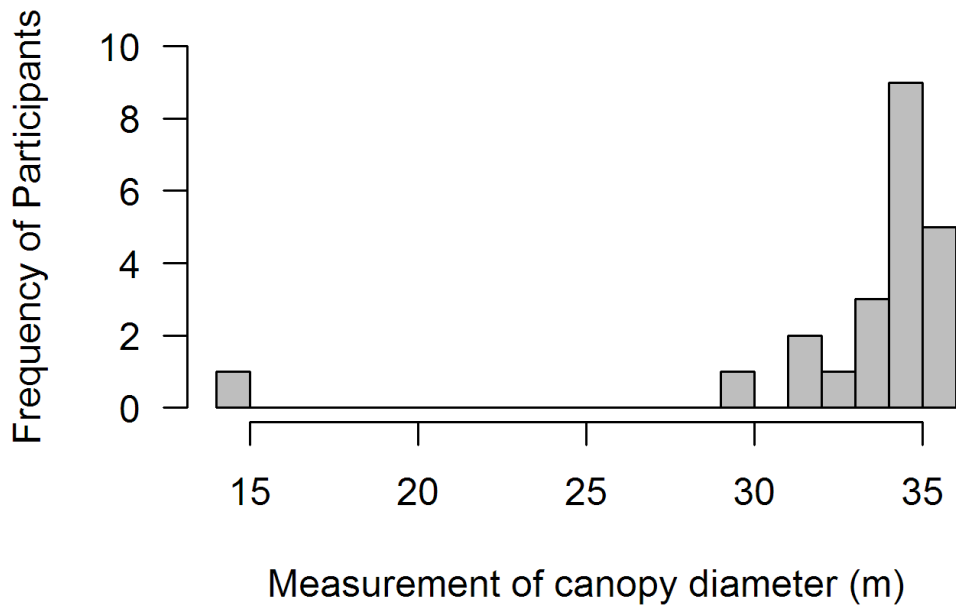
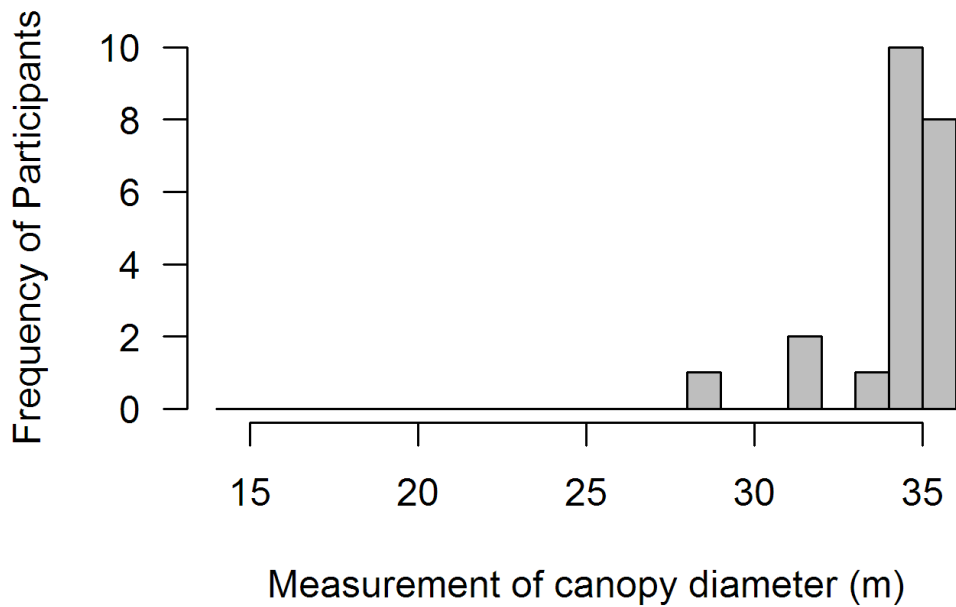
Scene B: 2D group (n=22)**Scene B: 3D group (n=22)**

Figure 5-15. Histograms of the Scene B longest canopy width measurements made by 2D group (top, n=22) and 3D group (bottom, n=22). Bin size = 1m.

Table 5-6. Distribution of measurements in the 2D and 3D groups for Scene B's canopy width.

| Distribution | 2D group (m) | 3D group (m) |
|--------------|--------------|--------------|
| Minimum | 8.853 | 25.91 |
| 1st Quartile | 32.32 | 33.15 |
| Median | 34.17 | 34.44 |
| Mean | 34.49 | 34.95 |
| 3rd Quartile | 35.53 | 35.53 |
| Maximum | 32.56 | 33.73 |

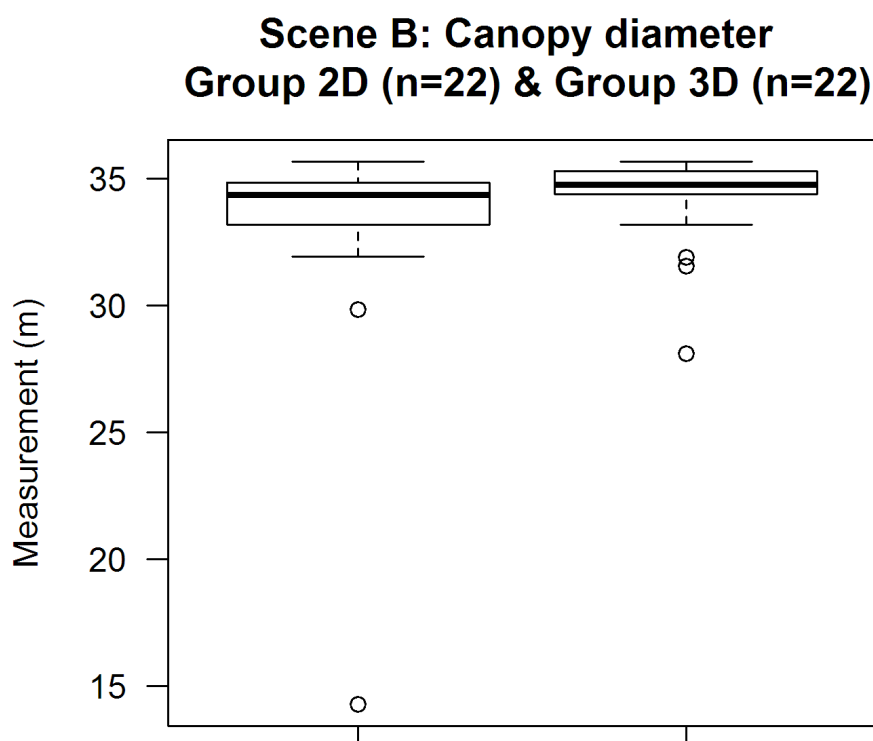


Figure 5-16. Boxplots of the Scene B longest canopy width measurements made by 2D group (left) and 3D group (right)

The Q-Q plots in Figure 5-17 illustrate the normality of the datasets, to help determine whether comparative statistical approaches should be parametric or not. The plots show the expected measurements (on the x axis), which assume a normal distribution (denoted by the solid line). In contrast to the equivalent planar plots in Figure 5-12, the expected results vary and are not a single value (i.e. 12.98m for planar task). Expected measurements range between approximately 32m and 36m for the 2D group and between approximately 33m to 36m for the 3D group. This implies that the answers from a manually measured volumetric feature are not expected to be consistent for both of the methods.

Comparison of observed measurements in both Figure 5-17 plots indicate a similar shape, whereby the interquartile range of observed samples (between -1 and 1 quantiles) loosely follow the normal distribution. Both groups have a maximum value of 35.53m, but their minimum values deviate from the norm to varying degrees. In the top plot of Figure 5-17 (2D group), one outlier measures 8.85m (at the -2 quantile) and the group's measurements, as a whole, have a range of 27.03m, whereas the 3D group results span 9.62m (min 25.91m). Here, the measurement of a volumetric point cloud with stereoscopic 3D visualisation yielded more precise results than those who carried out the 2D technique.

Since the p-values for the Shapiro-Wilk test is <0.05 (Table 5-7) the null hypothesis is rejected for both 2D and 3D. The implications of this normality test are that both datasets do not have a normal distribution; there is a difference between normal and observed values for each group. This means that parametric tests, which assume normality, cannot be applied. Instead, non-parametric or distribution-free tests are used when carrying out statistical analysis.

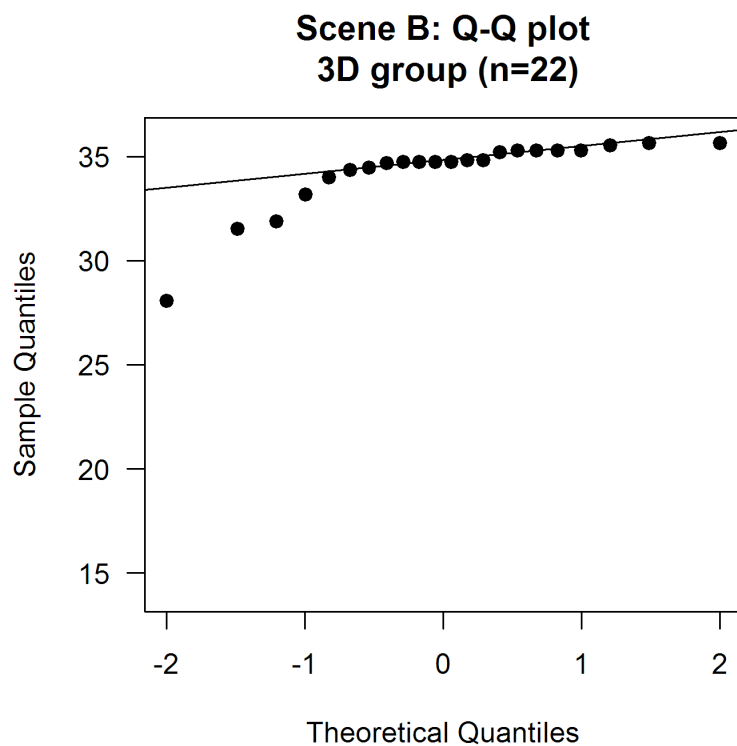
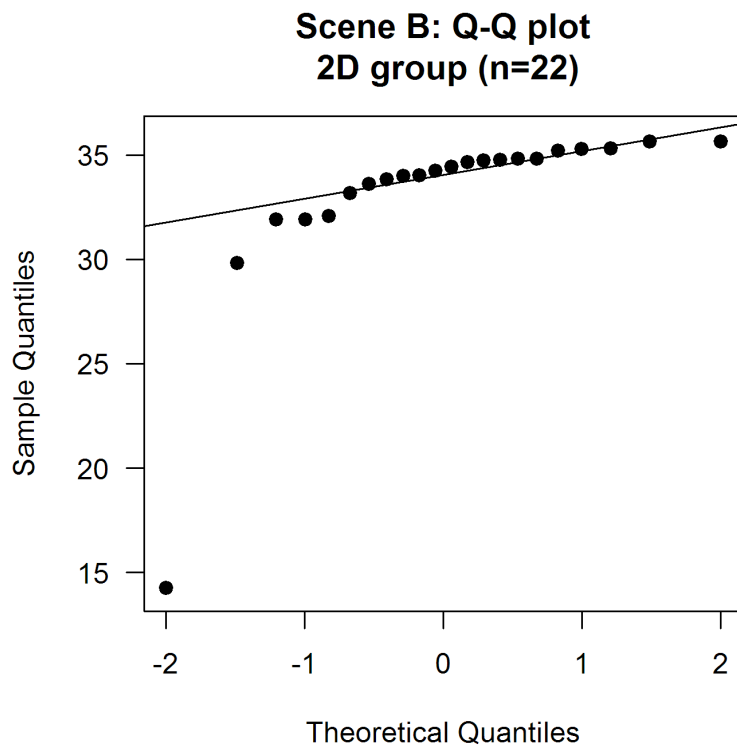


Figure 5-17. Theoretical vs. Sample quantile plots for Scene B 2D (top) and 3D (bottom). Along the x axis, 0 = mean (50% frequency) and deviations from this are one quartile (or 25%). Y axis = measurement results (in m) made by participants. One point represents a measurement made by one participant. Solid lines denote the theoretical normal distribution.

Table 5-7. Shapiro-Wilk normality test for 2D and 3D group, Scene B.

| Method | W value | P-value |
|---------------|----------------|----------------|
| 2D | 0.4953 | 1.187e-07 |
| 3D | 0.6868 | 1.337e-05 |

H₀ – No difference between actual and theoretical (p.value ≥ 0.05)

H₁ - Difference between actual and theoretical. (p.value <0.05)

Comparison of 2D vs. 3D volumetric feature measurements

Figure 5-18 plots the quantiles of the 2D and 3D participants against one and other. There is clustering of measurements leading up to and around 35m and there appears to be only one major outlier (measuring 8.85m) from the main cluster. There are no results higher than ~37m because there were no canopy points available beyond this for the participants to select. The Mann-Whitney test (Mann & Whitney, 1947) is used to compare the ranking of two non-normally distributed independent variables. Similarly to Scene A, an unpaired two-sample Wilcoxon signed rank test with continuity correction was performed in R (R Core Team, 2015), using the hypotheses below.

•H1.2b₀ – No difference between 2D and 3D measurements of canopy diameter (p.value ≥ 0.05)

•H1.2b₁ - Difference between 2D and 3D measurements of canopy diameter (p.value <0.05)

When comparing 2D and 3D groups using the Mann-Whitney test (U = 188.5, p-value = 0.2126), the estimated p-value is larger than 0.05, which indicates that the null hypotheses must be accepted. The means of the 2D and 3D samples for Scene B are not significantly different at the 0.05 confidence level, suggesting that they are from the same population, although two different techniques were used.

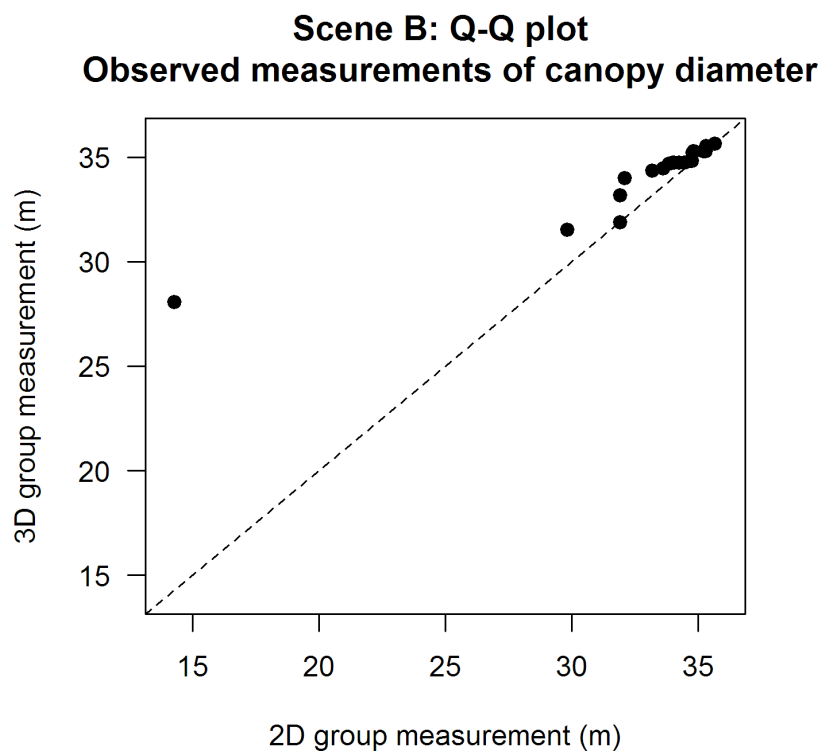


Figure 5-18. 2D vs. 3D quantile plots for Scene B, estimated longest canopy diameter. Dashed line represents 1.0 agreement.

Participant perception of 2D & 3D measurement

The frequency distribution of participants' ranked answers for Question 8 for Scene B, in Figure 5-19, exhibits a shape that is more Gaussian in nature, but the spread of confidence ranges between 2 and 7 (Scene A: 3-7), with the median peaking at the mid-way mark of 4 (Scene A: 6). This shows that participants were more comfortable with their measurements of the planar scene, compared to the irregular, volumetric canopy.

When discussing their measurement of Scene B, members of the 2D group, some of whose remarks are listed in Appendix K, felt it took more effort and precision than Scene A because it was not a regular structure and their own movement changed the view.

Q8: How accurately do you feel you measured Scene B?

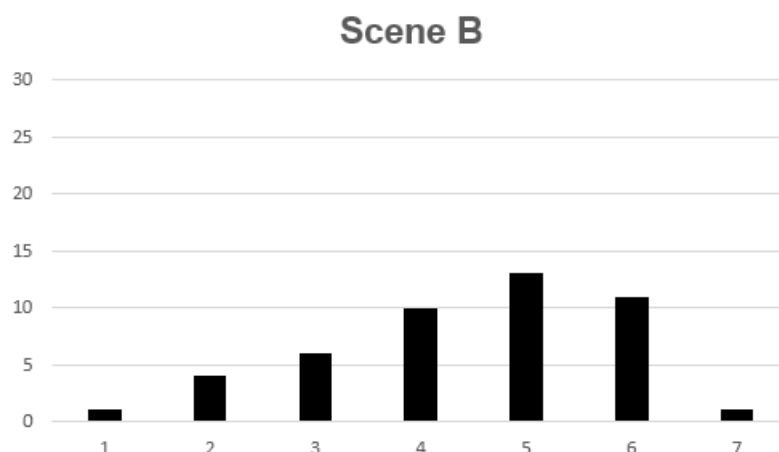


Figure 5-19. General participant responses to Question 8 – “How accurately do you feel you measured Scene B?”. Scores were ranked on a scale of 1-7, where 1 = Not at all accurately, 7 = Very accurately. Graph shows combined answers from 2D and 3D groups, n = 46 (inclusive of two participants whose quantitative results were discounted).

Participants commented that their movement, during exploration of the irregular feature, meant the ‘furthest point’ kept changing and this caused frustration during the task. Although a direct comparison cannot be made with the 3D group, this did not appear to be a complaint from those who had a stereoscopic view of the point cloud. Some users first attempted to measure the cloud from above, but were unable to differentiate between the ground and canopy points. This, coupled with the general irregularity of the shape, was highlighted by one of the participants:

“I’m not as happy as I was doing the other one – it’s much more difficult NOT in 3D. I think when you had better – the depth perception, then you had an idea distinguishing which [points] were ground and which were from a higher surface.

I think... because this is an irregular, or it seems to be, quite an irregular surface as well, it’s actually quite difficult to distinguish the points as well. So, I think those combined... just made life a bit more difficult.”

P47 (Scene B in 2D)

According to the 3D group of participants (Appendix K), the shape is not as straightforward and had no straight edges, which made it difficult to determine the longest axis. Among the remarks made by P25, it was reported that height could not be judged between the canopy and other points. The volunteer had a stereo score of 9/10, so it is surprising that they reported they were unable to judge height. This may be explained if the participant were at a distance from the data, when depth perception of objects is weaker. Although one participant mentioned they were unable to navigate during Scene B, it was the first scene that they came across and the random assignment of the Scene/Method orders takes this into account.

“Looking from above is a bit easier, but, again, only if you’re dead straight on to it. Um... and then you can’t be 100% certain whether that point’s... how high or how low it is in the canopy. So, you can’t judge height.”

P25 (Scene B in 3D) Plan View

5.3.3 Results summary

Results of Scene A and B measurement tasks reveal that measurement of planar and volumetric lengths from lidar point clouds are not significantly different between 2D *versus* 3D displays. Key quantitative findings for each scene are listed in Table 5-8 alongside the feedback provided by participants both during and after the tasks (Table 5-9).

Table 5-8. Summary of quantitative measurement task results.

| | measurement range | | |
|-------|-------------------|--------|------------------------|
| Scene | 2D (m) | 3D (m) | Significant difference |
| A | 0.91 | 1.84 | None |
| B | 27.03 | 9.62 | None |

Table 5-9. Summary of qualitative measurement task results.

| | participant comment summary | | |
|-------|---|--|---|
| Scene | point separation | straight edges | general 2D vs. 3D |
| A | In 2D, it was difficult to differentiate between ground and higher canopy points, from plan view. | Generally, straight edges in Scene A were apparent, but straight edges were especially apparent in 3D. | <ul style="list-style-type: none"> • The 2D display felt more familiar. • The 3D display gave a better sense of presence. • Participants that chose positions that actively exploited 3D technology had excellent stereo acuity. |
| B | | Lack of straight edges makes measurement of length more challenging. | |

5.4 Discussion

The findings of the measurement task are discussed in detail, in relation to Research Question 1 (section 2.4.1).

5.4.1 2D vs. 3D planar measurement precision (RQ1.2a)

Measurements of the planar roof length from the 2D group (n = 21) and the 3D group of participants (n = 22) were compared to answer RQ1.2, restated below. The result is consistent with the null hypothesis – there is no significant difference between the distributions of the 2D and 3D-derived measurements. These results are based on the distribution of measurements.

RQ1.2 Is there a significant difference between point cloud measurements made in 2D, in comparison to those made in 3D for **(a) a planar feature?**

The similarities in distributions may be explained by psychological theory. In both 2D and 3D, humans can perceive an interpolated surface in between a

continuous surface of textures and this is not affected by the density of the textures (Wilcox & Duke, 2005). This implies that any abrupt changes in a smooth planar surface, such as the roof in Scene A, can be easily detected, e.g. edges or disruptions in the plane. This human ability to cognitively form surfaces would have been experienced by both 2D and 3D participants.

Both 2D and 3D groups of participants felt that the planar measurement was 'straightforward'. The distinct, straight edges of the house roof made it easy to determine the measurement and identify its start and end points. It appeared, from comments, that the participants from the 3D group considered the straight edges especially noticeable. Despite this observation, the results imply that 2D representation of straight edges may be sufficient for linear measurements. The volunteers also stated their intentions to position themselves perpendicular to the measurement and view the extents of the feature.

5.4.2 2D vs. 3D volumetric measurement precision (RQ1.2b)

Scene B provided a volumetric feature for analysis, in contrast to Scene A, to address the objective below. Similarly to the planar feature, there was no significant difference between 2D and 3D measurement values generated from Scene B.

| |
|--|
| <p>RQ1.2 Is there a <u>significant difference</u> between point cloud measurements made in 2D, in comparison to those made in 3D (b) a volumetric feature?</p> |
|--|

The difference between 2D and 3D distributions of measurement were not significant. Comments in Appendix K suggest they were not confident and they felt that the task demanded more precision during point selection. Other qualitative results showed all participants (2D and 3D groups) found the volumetric measurement task harder than the planar measurement. Those who carried out the Scene B measurement in 3D remarked that the scene itself made the task hard to work out. The volumetric nature of the point cloud,

where the points fill a 3D space, means there is occlusion and the scattered spatial arrangement of the points meant that movement was required.

The irregularity of the volumetric feature was more challenging, more so for members of the 2D group, who reported that they were unable to differentiate between points of different depths. When anticipating their measurement actions, participants stated their intentions to look from above, but 2D participants could not differentiate the ground and others (P47 and P46). This may indicate a need for users to remove points from view in 2D, to alleviate the sub-task of point differentiation before carrying out any decision-making. This is an option available in current 2D software (Terrasolid, 2013a; Girardeau-Montaut, 2014; McGaughey, 2014) and suggests that HCI capabilities already provide the necessary work-around for the visual perception drawbacks of 2D manual measurement.

Although participants' ability to visually interpolate surfaces could be a contributing factor to similar 2D and 3D measurement of planar features, the vegetation in Scene B had a volumetric arrangement of points. It was assumed that when the surface has a higher dimensionality (Jones *et al.*, 2008), whereby the points are distributed throughout a volume, a stereoscopic viewing method would help the viewer to detect disparity in depth. However, the use of binocular 3D depth when viewing a complex 3D volumes puts more load on the binocular visual system. From a human perception perspective, a tree is a complex volume (Harris, 2014). In her cognitive psychology study (Harris, 2014), users were presented with different objects at different depths and orientations, to assess the thickness of a volume. This is comparable to this study's longest diameter concept. Although the spherical points provided by the lidar scans are not themselves oriented, the larger patterns and surfaces that they form (clusters of leaves, branches), as Wilcox and Duke (2005) suggest, will be positioned and oriented at a range of depths. Harris found that for a complex volume, binocular (3D) vision could not perceive depth of all of the elements at once. Initially it was expected that the outputs of the 3D visual task would outweigh the lower cognitive effort to view a flattened representation of the data that 2D provides. In McIntire *et al.*'s (2014) review of 184 2D vs. 3D experiments, 25% of the comparisons found that stereoscopic 3D visualisations

did not add value to a range of tasks. It was identified that simple tasks and those that are not reliant on depth information did not benefit from additional stereo depth (McIntire *et al.*, 2014). In the volumetric scene, it was expected that the measurement would be easier because of the added depth, but the linear measurement task itself, not the point cloud structure, could be too simple to require added binocular depth. In addition to this, for Scene A, the planar feature may have created linear perspective depth cues that the participants could take advantage of in both 2D and 3D.

The shortest diameter that was measured of Scene B, at 8.85m by a 2D participant and this skewed the 2D results. Their measurement could be a result of the lack of depth associated with the 2D display of the points; the participant was not able to judge the appropriate depth. However, the nature of the human-centric study means that judgement can be affected by uncontrolled conditions, including emotional factors (Eysenck, 2001) and attention. These were not measured during the experiment, so the reason for the outlier is unclear. For the measurement of Scene B, the volunteers of both methods found the relative complexity of the feature challenging compared to the planar feature in Scene A.

The overall results indicate that 3D visualisation does not add value to linear measurement of point clouds and 2D methods. However, the findings contribute to the evaluation approach for applied psychology methods. Reflections on the approach used in this study are discussed in the next section.

5.5 Reflection on methodological development (RQ3)

This research experiment is novel and aspects of the measurement task approach are considered in relation to impact on the results. This section

partially addresses the RQ3, restated below, which is also posed in Chapters 6 and 7 in relation to other aspects of the study.

RQ3 How effective is the methodology at comparing 2D vs. 3D visualisation of lidar point clouds?

In addition, recommendations are made for consideration during the different stages of interaction design evaluation, as outlined by Preece *et al.* (2007).

5.5.1 Interaction technique

During measurement of both planar and volumetric features, users were trained to select points. This data interaction was purely to register the points that participants estimated to represent the start and end of the length that they were measuring. The method did not allow sophisticated interaction with the data, partly because the task was fulfilled by point selection and partly because the pilot revealed that participants would be cognitively overloaded by the number of new concepts (lidar-related) and manual tasks (gamepad use). This led to the adoption of basic data-handling technique. After reviewing 184 2D vs. stereo 3D human performance experiments, McIntire *et al.* (2014) concluded that one of the most useful characteristics of the 3D technique was object manipulation (McIntire *et al.*, 2014). This implies that the basic interaction processes used here for feature measurement potentially sold short 3D visualisation.

One participant felt that a mouse would be better for point selection during the Measurement Task. However, the use of 2D device, *i.e.* manoeuvred on a 2D plane, for a 3D space would introduce issues. Within a geoscience visualisation created by Kreylos *et al.* (2006), users were able to directly manipulate a virtual 3D lidar dataset with one hand and draw vectors with the other. An example of hands-on manual interaction with data was proposed by Trinks *et al.* (2005), through photorealistic virtual analysis of geological structures. Users are able to select points, save them as objects, and from them extrude spline layers to help determine rock formation. Despite their more involved, non-2D data interaction techniques, these virtual lidar studies did not make a direct comparison to an

equivalent 2D visualisation, so it is not clear whether their data manipulation methods are the same in 2D or if they are significantly more accurate. Jáuregui *et al.* (2012) describe the combination of 3D cursors (e.g. an avatar hand, which appears in front of the user (Ware, 2012) with 2D device for selecting data in a virtual environment. Although there was an improvement in depth perception when using a 3D cursor, instead of a 2D cursor, user selection performance decreased. The developers of more recent HMD hardware recognise the HCI challenge posed by virtual cursors, a seemingly simple GUI design (Edge®, 2013). The use of 2D crosshair in a 3D projection would affect stereovision because its on-screen position would interfere with nearby 3D objects (Hoffman *et al.*, 2008). Although this was the same in 2D and 3D, the effect in 3D could be a disadvantage to those stereo users. On a stereoscopic display, the focus of the eyes is constant, on the physical screen, although the objects may appear in front or behind the screen. However, in reality, when we focus on something, other objects that are further away from or closer to us blur as our eyes converge or diverge. Edler (2014) uses a static lenticular 3D projection, which shows images as increasingly blurred with increasing (positive or negative) distance from the projection layer, although the view is restricted to one perspective and resolution. In contrast, in this 2D vs. 3D study, participants were able to change the perspective and level of detail, which in turn changed the distance of objects from projection screen, where the eyes were focussed. To apply a depth-of-field (DoF) blurring, according to the participant's gaze, eye-tracking information could be fed back to the visualisation system, blurring the points on which the eyes are not fixated. Hillaire *et al.* (2008) studied first-person navigation in a virtual environment and participants' felt that dynamic DoF blur, based on eye-tracking, improved the visual effects. The addition of this visual effect was out of the scope of this study, but a more dynamic effect could improve the 2D and 3D visualisation conditions in future studies.

In a comparative study, there is a need to make the visualisation methods as similar as possible so that they can be fairly evaluated against one and other. This study used the same gamepad device for both 2D and 3D visualisations to standardise the HCI. However, another input device may be more appropriate

for each of the visualisation approaches. For example, instead of a gamepad, whose configuration is controlled by joysticks, a 2D visualisation may benefit from a 2D input device (e.g. mouse). A 3D visualisation approach might be enhanced by an interaction device that uses 6 DOF motion control in 3D space, such as the Leap Motion (www.leapmotion.com) or yet-to-be-released Oculus Touch (www.oculus.com). In line with Mark's (1992) spatial data handling recommendation, three-dimensional interaction technique would be cognitively appropriate for lidar measurement, although would require user training. Banic (2014) explores 3D interaction techniques by using a physical, real-world installation to study user interaction with volumetric data, instead of observing users interact with virtual data. Hand movements of the ten participants were tracked in a real space as they explored the points, which were cotton wool balls suspended on string. Banic (2014) strips back the complex challenges of devices and interfaces to put human interaction needs first. This physical observation technique would also eliminate restrictions cause by the use of navigation devices, which participants may find intimidating. Chen *et al.* (2013) provide an extensive review of depth sensors, particularly Microsoft's Kinect (Microsoft Corporation, 2015) and predict that interaction sensors will be incorporated into desktop and mobile devices. This suggests that more intuitive HCI technologies will become more commonplace and allow humans to interact with 3D data within 3D space. This is in contrast to our current familiarity with a 2D mouse, which does not provide a metaphorical representation of our true interactions in space (Mark, 1992). The advancements in the remote detection of human motion with an increased understanding of HCI allow for the adoption of more natural handling of datasets, whether the task is measurement or data manipulation.

An increasing distance of the user from the points may decrease the influence of stereoscopic depth perception on the 3D user, meaning that the stereo depth cues could be similar to 2D (*i.e.* non-existent) when at a certain proximity from the data. Stereo depth from binocular disparity ceases at 215m (Howard & Rogers, 2012, p. 327) for someone with 60 secs of arc (between 6-7/10 on the Randot test (Stereo Optical Co., 2009)), but Piryankova *et al.* (2013) claim that stereoscopic projection only affects near distance (2m) estimates in a virtual

environment. This is also influenced by oculomotor responses, see Howard and Rogers (2012) for detail and further reading, but literature suggests that user distance from the virtual 3D objects could affect stereoacuity. Despite this possibility, here, the measurement interaction design required the participants to place the point within the centre of the crosshair target, which put them into close contact with their target point. For the 3D users, the sensation of depth would have been strong because of the proximity to the point.

Method Reflection 1: *The results of the measurement task may have been affected by the interaction design.*

5.5.2 Vegetation Analysis Approach

The participant instructions for Scenes A and B are different because of the shape and dimensionality of the features. With the planar, rooftop (Scene A), it is quite clear which length the participants must estimate, their measurement decision depends on the presence of points, but the canopy in Scene B requires a subjective evaluation of the shape before selecting the relevant points. This additional analysis step is a pre-requisite for the canopy measurement and this extra cognitive requirement could influence the participants' feedback to Question 8, regarding perceived accuracy of measurement. However, in their own crown analysis study, in which they used Pointools® View Pro software (now Bentley Pointools v8i, www.bentley.com) for virtual (still 2D) measurements, Yang *et al.* (2013) acknowledge that crown-derived parameters are subjective. This, coupled with the increased length of the feature, is reflected in the increase in histogram bin size between planar and volumetric measurements (from 0.1m to 1m).

The participants did not have any tools to assist in their measurement of the diameter, so this measurement could be used for validation. Although the pictorial instructions indicated this, the participants were not explicitly told to measure horizontally. Furthermore, the participants could only select points, but the nature of the point cloud meant that there may not have been points along the same xy plane, *i.e.* there may not have been points available.

Method Reflection 2: *Vegetation canopy measurement in 2D and 3D may have benefitted from tools.*

5.5.3 2D vs. 3D Assessment Methods

Alternative performance indicators could help to identify other 2D vs. 3D differences when none is observed in precision alone, as was the case in this study. In the McIntire *et al.* (2014) review of 162 2D vs. 3D studies, completion time was noted to common paired with another performance indicator. Although the pilot participants felt that scrutiny of their performance completion time affected their confidence during tasks (section 3.3.3), the coupling of more than one performance measure, such as the inclusion of time, would help to determine more conclusive results regarding the effects of 2D vs. 3D.

Method Reflection 3: *The measurement task results were not significantly different, but other performance indicators could reveal differences*

5.6 Summary

The outcome of the Measurement Tasks revealed that the general distributions of 2D and 3D estimations were not significantly different when measuring (a) the length of a roof edge in 3D and (b) the longest diameter of a volumetric tree canopy. The reason for the similarities of the planar measurement could be because straight edges are apparent (already) in 2D, which made it easy in both tasks. The irregularity of the volumetric feature made it difficult for the participants to estimate the more subjective length. Overall, the 2D visualisation may have provided sufficient depth cues, whose enhancement by binocular 3D vision did not add value to the task. The analysis of additional factors may reveal significant differences between 2D and 3D performance and recommendations for further research are made in the conclusion.

6.

Interpretation

6. Interpretation

6.1 Introduction

The interpretation task investigates human interpretation of lidar point clouds, under 2D and 3D visualisation conditions:

RQ2 Is lidar point cloud feature interpretation more accurate when carried out via 2D (monocular) or 3D (binocular) visualisation?

The task was carried out by participants to explore any differences in the accuracy of 2D and 3D object recognition from a lidar point cloud. This chapter explains the further methods required to develop the task, presents results, and discusses the study's findings in the context of the literature. Further to RQ2, its sub-questions, below, ask (RQ2.1) how does feature geometry affect the results and is there a difference between 2D and 3D visual interpretation of the selected features? Concerning their wider acquisition environment (RQ2.2), does its complexity affect the accuracy of 2D vs. 3D visual interpretation results? Broadly speaking, the interpretation task assesses how 2D interpretation of lidar point clouds compares to interpretation via a proposed stereoscopic 3D display.

RQ2.1 For point cloud features of varying geometries,
 (a) what is the 2D and 3D interpretation accuracy? (b) is there a significant difference between interpretations of the point cloud features made in 2D, in comparison to those made in 3D?

RQ2.2 Does the complexity of the lidar point cloud acquisition environment affect the accuracy of 2D vs. 3D visual interpretation results?

(a) On flat terrain (b) On sloped terrain

6.2 Method

The interpretation-specific methodologies are explained in this section and add-on to the general experiment method (Chapter 3).

6.2.1 Task development

Scene Selection

The AOIs that were chosen for the interpretation task, Scenes C and D, had to fulfil the following criteria:

- Heterogeneous coverage e.g. vegetation, building, other features.
- Point clouds that have regions of points that are classified incorrectly by automated processing.
- Study sites with different terrain conditions (e.g. flat, inclined), following the example of the Meng *et al.* (2010) who investigated the effect of auto-classification algorithms on 15 different AOIs.

A mixed coverage of features provides the participants with a range of objects to classify. The underlying cause of misclassified points could be, for example, vegetation in close proximity of building. The sloping versus flat environments allow the study to explore whether the application of depth perception is affected by the geographical environment.

Selection of points of interest (POIs)

Points of interest (POIs) were picked via a standard 2D computer interface, using open-source lidar software CloudCompare (Girardeau-Montaut, 2014). Five POIs were selected for each AOI and are shown on a plan view image of each AOI, Figure 1-1 6-1 and Figure 6-2.

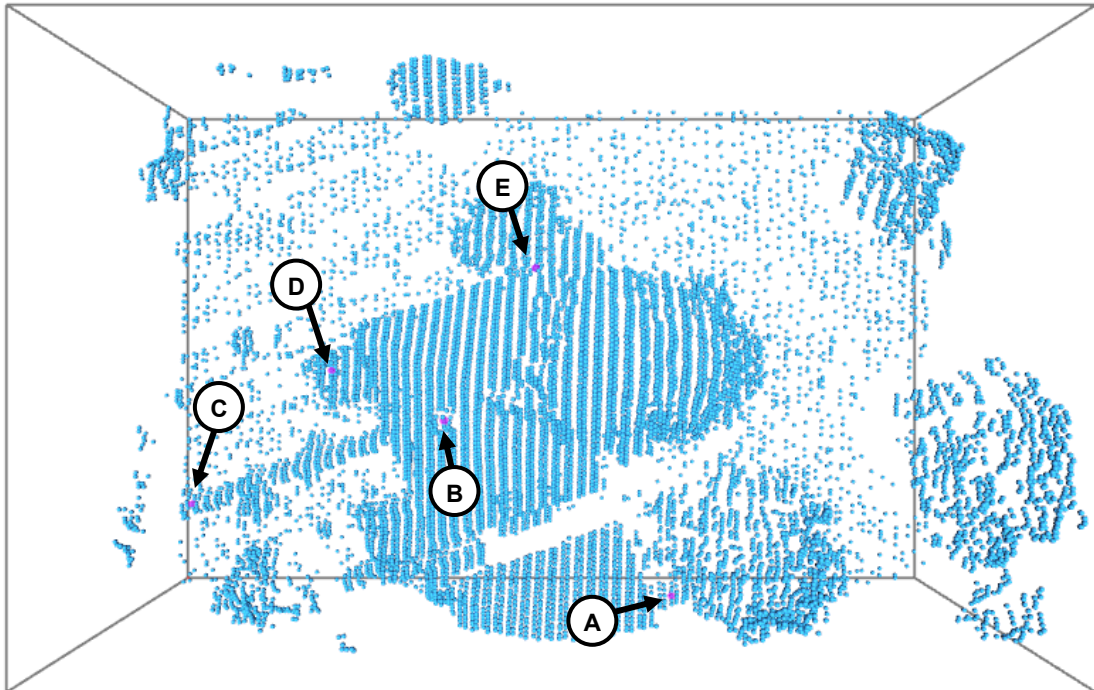


Figure 6-1. Plan view of Scene C point cloud, showing the location of the points of interest that were interpreted by the participant. POIs are highlighted by flashing pink glow. Lidar data © Airbus Defence and Space Ltd. (2013).

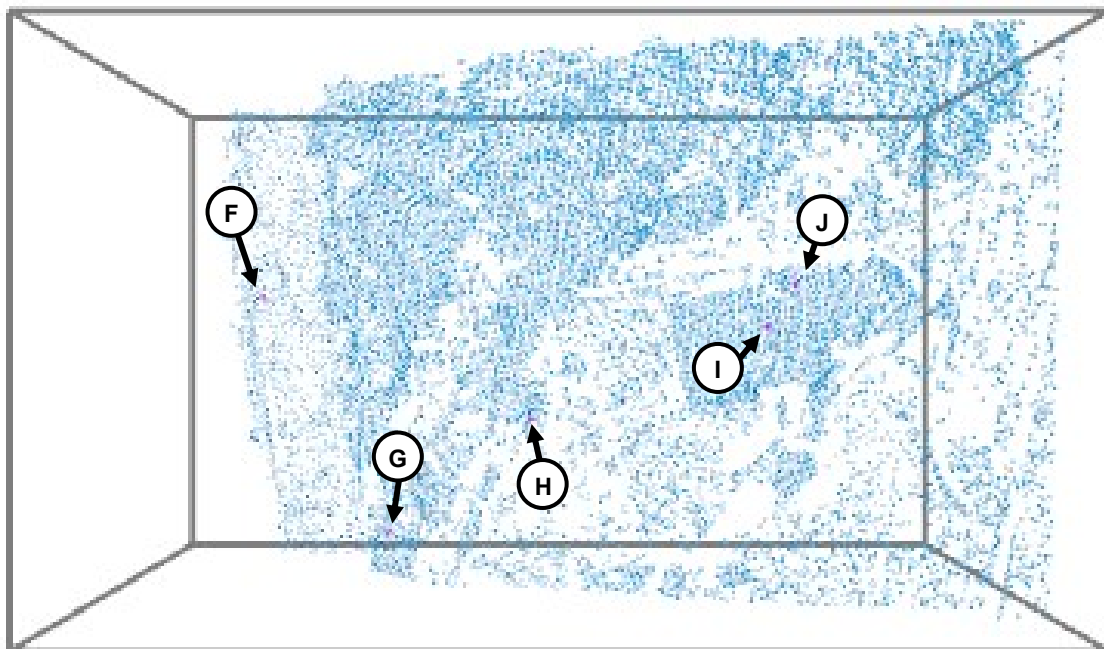


Figure 6-2. Plan view of Scene D point cloud, showing the location of the points of interest that were interpreted by the participant. POIs are highlighted by flashing pink glow. Lidar data © Airbus Defence and Space Ltd. (2013)

The POIs were selected based on the point cloud structures to which they belong, from the obvious (road or distinct tree) to the geometrically complex (unusual forms and ambiguous boundaries between features). These complex areas were identified from areas of misclassification of the classified lidar point cloud, using processes explained in Chapter 3. An initial feature validation technique involved lidar interpretation and a comparison of the POI positions against orthophotographs (GeoPerspectives, 2013b; GeoPerspectives, 2013a) and GoogleEarth StreetView photography (Google, 2014).

Validation of each POI

The validation of the ten POI features was strengthened through discussion with an experienced lidar processing manager at Airbus Defence and Space (formerly Astrium UK). The lidar data were used for validation by the expert operator, but orthophotographs and Google Streetview images are presented here for context, a list of their acquisition dates are shown in Table 6-1.

The visualisation technique used for the interpreter's visual analysis was via a standard 2D desktop screen, with plan and cross-sectional point cloud views in Terrasolid (Terrasolid, 2013a). Interpretation commentary is summarised alongside aerial imagery and point cloud cross-sections that were viewed during the visual check, found in Table 6-2 to Table 6-7. The imagery in the tables has no scale, so it is used for illustration purposes. The large white arrows containing the text, 'st. view', indicate the approximate direction from which the Google StreetView imagery (Google, 2014) was captured. It was evident from the lidar point cloud that one point in Scene C is the roof of an extension (D) and POIs in Scene D were located on the road (F) and on a tree (H). For this reason, these POIs do not feature in the tables.

Table 6-1. Summary of reference answers based on lidar acquisition interpretation by one expert operator. Data acquisition dates are listed for lidar, orthophotos and Google StreetView imagery.

| Scene | POI | Reference Interpretation | Lidar acquisition date | Ortho-photo date | Google StreetView date |
|-------|-----|--------------------------|------------------------|------------------|------------------------|
| C | A | Vegetation (shrub) | 25-Nov-08 | 14-Apr-07 | June-08 |
| | B | Chimney | 25-Nov-08 | 14-Apr-07 | June-08 |
| | C | Hedge | 25-Nov-08 | 14-Apr-07 | June-08 |
| | D | Extension | 25-Nov-08 | 14-Apr-07 | - |
| | E | Tree | 25-Nov-08 | 14-Apr-07 | June-08 |
| D | F | Road | 04-Jul-08 | 27-Jun-10 | - |
| | G | Cliff | 04-Jul-08 | 27-Jun-10 | July-08 |
| | H | Vegetation (tree) | 04-Jul-08 | 27-Jun-10 | - |
| | I | Rooftop | 04-Jul-08 | 27-Jun-10 | Sept-08 |
| | J | Rooftop | 04-Jul-08 | 27-Jun-10 | Sept-08 |

Table 6-2. Feature verification of Point of Interest (POI) A in Scene C. The white dot (just seen) in the orthophoto shows the POI location. The $A_1 - A_2$ line in the ortho data is shown as 2D transect cross-section in the lidar data. The POI is located within the white circle. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophoto © GeoPerspectives (2013). Un-scaled screenshot images taken from Terrasolid lidar software GUI (Terrasolid, 2013b).

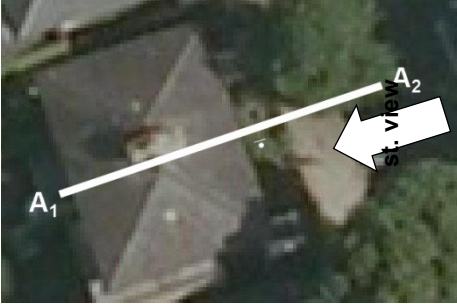
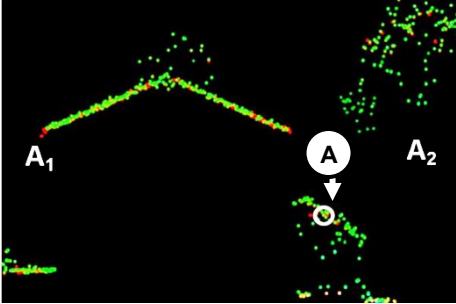

| POI A | Reference data used for feature validation |
|--------------|--|
| Data | <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;">  <p data-bbox="432 853 839 882">Plan view (N-orientated, with lean)</p> </div> <div style="text-align: center;">  <p data-bbox="938 853 1329 882">$A_1 - A_2$ point cloud cross-section</p> </div> </div> <div style="text-align: center;">  <p data-bbox="432 1496 1297 1547">W-facing, taken June 2008 (Google, 2014) https://www.google.co.uk/maps/@51.5067468,-0.2124308,3a,75y,266.52h,84.88t/data=!3m6!1e1!3m4!1svRPcNZveL9vPUWXWLDvEQI2e0!711331218i6656?hl=en-GB</p> </div> |
| Expert Check | <p>From looking at the image, my first thought is a tree in the front garden. From looking at the lidar (checking multiple cross-sectional views) it is not part of the main house, could be part of an extension. The shape appears man-made, but from the data, you can see that it's green foliage, so more 'foliage'.</p> |
| Feature | Vegetation → small tree / shrub |

Table 6-3. Feature verification of Point of Interest (POI) B in Scene C. The white dot (just seen) in the orthophoto shows the POI location. The $B_1 - B_2$ line in the ortho data is shown as 2D transect cross-section in the lidar data. The POI is located within the white circle. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophoto © GeoPerspectives (2013). Un-scaled screenshot images taken from Terrasolid lidar software GUI (Terrasolid, 2013b).

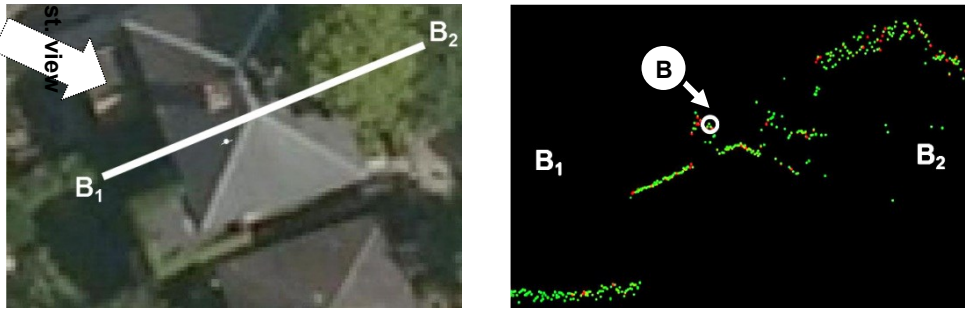

| POI B | Reference data used for feature validation |
|--------------|--|
| Data |  |
| | <p>Plan view (North-orientated, with lean) $B_1 - B_2$ point cloud cross-section</p> |
| |  |
| | <p>SE-facing, taken June 2008 (Google, 2014) https://www.google.co.uk/maps/@51.5069375,-0.2131515,3a,75y,129.72h,88.52t/data=!3m6!1e1!3m4!1s0rSXhBsVeFbYoQQKnXLhdAl2e0l7i13312!8i6656?hl=en-GB</p> |
| Expert Check | Chimney. An aerial would be only 1 or 2 points. |
| Feature | Chimney |

Table 6-4. Feature verification of Point of Interest (POI) C in Scene C. The white dot (just seen) in the orthophoto shows the POI location. The C₁ – C₂ line in the ortho data is shown as 2D transect cross-section in the lidar data. The POI is located within the white circle. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophoto © GeoPerspectives (2013). Un-scaled screenshot images taken from Terrasolid lidar software GUI (Terrasolid, 2013b).

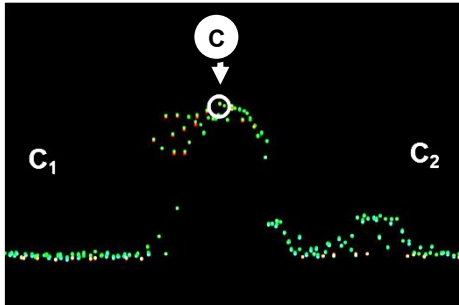
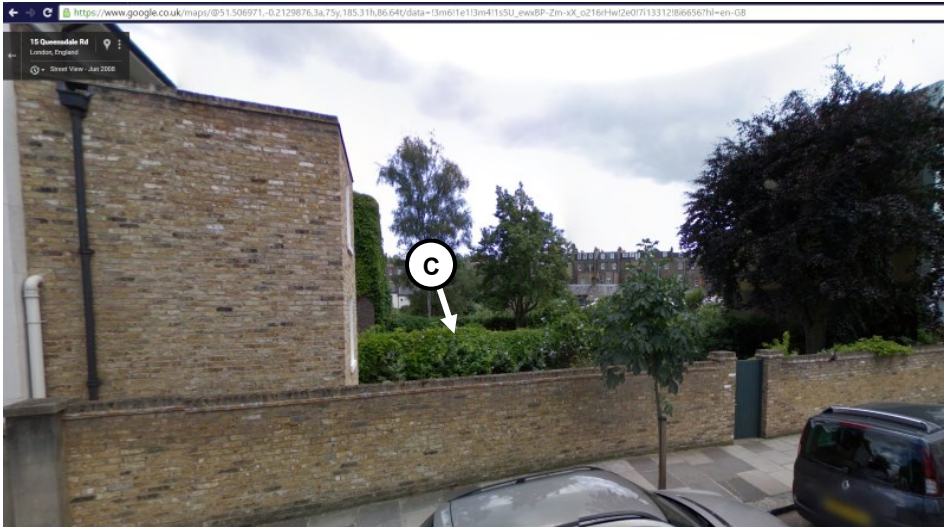
| POI C | Reference data used for feature validation |
|--------------|---|
| Data | <div style="display: flex; justify-content: space-around;"> <div data-bbox="427 517 887 871">  <p data-bbox="427 902 818 965">Plan view (North-orientated, with lean)</p> </div> <div data-bbox="922 568 1382 871">  <p data-bbox="922 902 1315 931">C₁ – C₂ point cloud cross-section</p> </div> </div> <div data-bbox="427 994 1382 1518">  <p data-bbox="427 1550 1347 1599">S-facing, image taken June 2008 (Google, 2014) https://www.google.co.uk/maps/@51.506971,-0.2129876,3a,75v,185.31h,86.64t/data=!3m6!1e1!3m4!1s5U_ewxBP-Zm-xX_o216rHw!2e0!7i13312!8i6656?hl=en-GB</p> </div> |
| Expert Check | <p>If no imagery, it might look like a wall, but it's rounded on top, which suggests vegetation. Point C feature cross-section shows one side is well-trimmed, but the other is more overgrown.</p> |
| Feature | Vegetation → hedgerow |

Table 6-5. Feature verification of Point of Interest (POI) E in Scene C. The white dot (just seen) in the orthophoto shows the POI location. The $E_1 - E_2$ line in the ortho data is shown as 2D transect cross-section in the lidar data. The POI is located within the white circle. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophoto © GeoPerspectives (2013). Un-scaled screenshot images taken from Terrasolid lidar software GUI (Terrasolid, 2013b).

| POI E | Reference data used for feature validation |
|--------------|---|
| Data | <div style="display: flex; justify-content: space-around;"> <div data-bbox="427 495 887 869">  <p data-bbox="427 898 820 958">Plan view (North-orientated, with lean)</p> </div> <div data-bbox="922 566 1382 869">  <p data-bbox="922 898 1315 927">$E_1 - E_2$ point cloud cross-section</p> </div> </div> <div data-bbox="427 992 1382 1507">  <p data-bbox="427 1541 1302 1585">SW-facing, taken June 2008 (Google, 2014) https://www.google.co.uk/maps/@51.5069723,-0.2125533,3a,75y,252.96h,91.07t/data=!3m6!1e1!3m4!1s8k0NbPycP3ZhpzFw8mwQ!2e0!7i13312!8i6656?hl=en-GB</p> </div> |
| Expert Check | <p>From a cross-sectional check, it could be both [house or tree]. Looks in line with the roofline. Same height as the roof. From checking against the ortho, it's just beyond the roofline. If you look along the rest of the roof and follow the roofline in cross-section, it is vegetation, just!</p> |
| Feature | Vegetation → tree |

Table 6-6. Feature verification of Point of Interest (POI) G in Scene D. The white dot (just seen) in the orthophoto shows the POI location. The $G_1 - G_2$ line in the ortho data is shown as 2D transect cross-section in the lidar data. The POI is located within the white circle. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophoto © GeoPerspectives (2013). Un-scaled screenshot images taken from Terrasolid lidar software GUI (Terrasolid, 2013b).


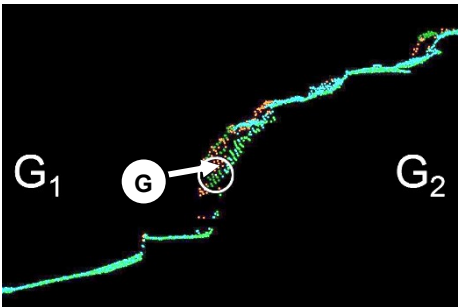

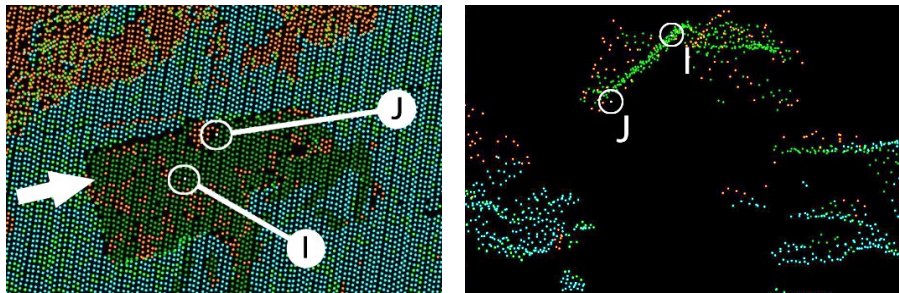
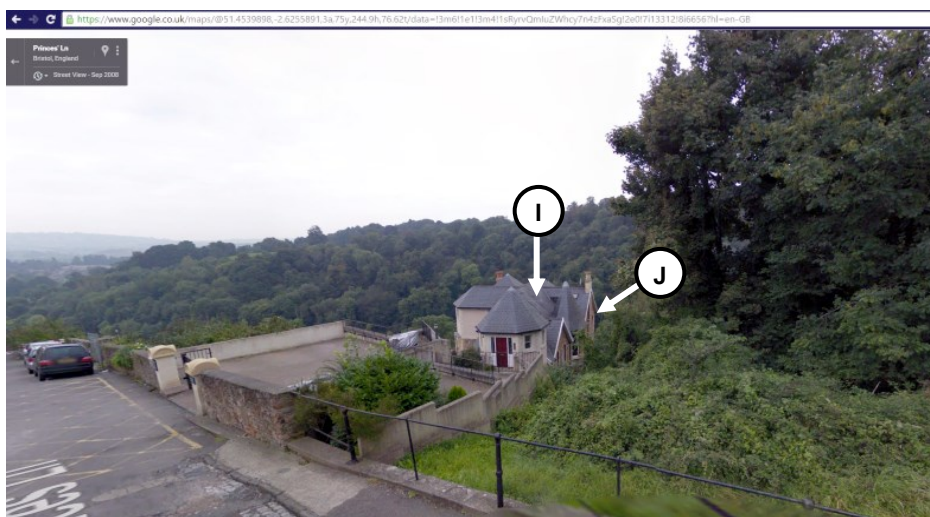
| POI G | Reference data used for feature validation |
|--------------|---|
| Data | <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;">  <p data-bbox="427 853 711 909">true-colour point cloud (Geostore & Terrasolid)</p> </div> <div style="text-align: center;">  <p data-bbox="924 853 1321 882">$G_1 - G_2$ point cloud cross-section</p> </div> </div> <div style="text-align: center; margin-top: 10px;">  <p data-bbox="427 1507 1359 1554">StreetView image from June 2008 (Google, 2014) https://www.google.co.uk/maps/@51.453536,-2.6267006,3a,75y,54.5h,131.62t/data=!3m6!1e1!3m4!1sdL_VUv9rseZSD2a-asoNeMq!2e0!7!13312!8i6656?hl=en-GB</p> </div> |
| Expert Check | <p>Without classification knowledge/imagery, it doesn't look like building [as categorised by auto-classification]. Google StreetView shows it is next to a retaining wall, which, in commercial lidar terms, would be considered ground. A cross-section of the lidar [right] also reveals that the point itself is on the cliff-face next to the retaining wall. It is ground on a vertical/cliff-face.</p> |
| Feature | Cliff-face/ground |

Table 6-7. Feature verification of Point of Interests (POI) I and J in Scene D. The white dot (just seen) in the orthophoto shows the POI location. The POIs is located within the white circle. Lidar data © Airbus Defence and Space Ltd. (2013). Orthophoto © GeoPerspectives (2013). Un-scaled screenshot images taken from Terrasolid lidar software GUI (Terrasolid, 2013b).

| POI I & J | Reference data used for feature validation |
|--------------|--|
| Data |  |
| | <p>Left: point cloud in plan-view (N-orientated) with white arrow indicating the direction of view of the cross-sectional view (right). Here, green = building, orange = veg, blue = ground.</p> |
| |  |
| | <p>SW-facing, taken from road above the AOI in Sept 2008 (Google, 2014) https://www.google.co.uk/maps/@51.4539894,-2.6255891,3a,75y,244.9h,76.62t/data=!3m7!1e1!3m5!1sRyrvQmluZWlucy7n4zFxaSg!2e0!5s20080901T00000!7i13312!8i6656?hl=en-GB</p> |
| Expert Check | <p>Re. rough nature of the points, there is always some scatter, could be to do with the surface material.</p> |
| Feature | <p>Building</p> |

6.2.2 Experiment instructions

During the experiment, the interpretation task (Task 2) was preceded by the measurement task. Although the volunteers were shown ground reference data once they had measured Scenes A and B, these media were not available to the participants when viewing Scenes C and D (Figures 6-4 and 6-5). The only auxiliary information provided was a bounding box indicated the axes of the datasets and a 5m³ cube that was shown by the researcher, for scale reference.

Familiarisation of scenes – grid view

This stage of the experiment took place prior to the POI interpretation, to allow the volunteers to get a sense of the scene they were about to interpret. The 6-cell grids (Appendix H) were laid over the displayed AOI (C or D) prior to exploration of the scene. Participants were informed that there will be a brief discussion about what features they think might be present in each grid cell, if any, and were given 2-3 minutes, as required, to navigate and familiarise themselves with the scene. The participant described the contents of each of the 6 grid cells from plan view.

Flashing points (POI interpretation)

Participants were instructed to visit each of the flashing points, in any order, and describe the feature to which the point belongs, a tree, a house, *etc.* They were each encouraged to explore the scene to help figure out the POI features. Figure 6-3 details the set-up during the task, showing an examples of the participant's data interaction during the task. Since the classifications were described verbally, the interviewer carried out translations or asked for verification when required. The outputs of this task were audio recordings and notes made by the interviewer during the trial.

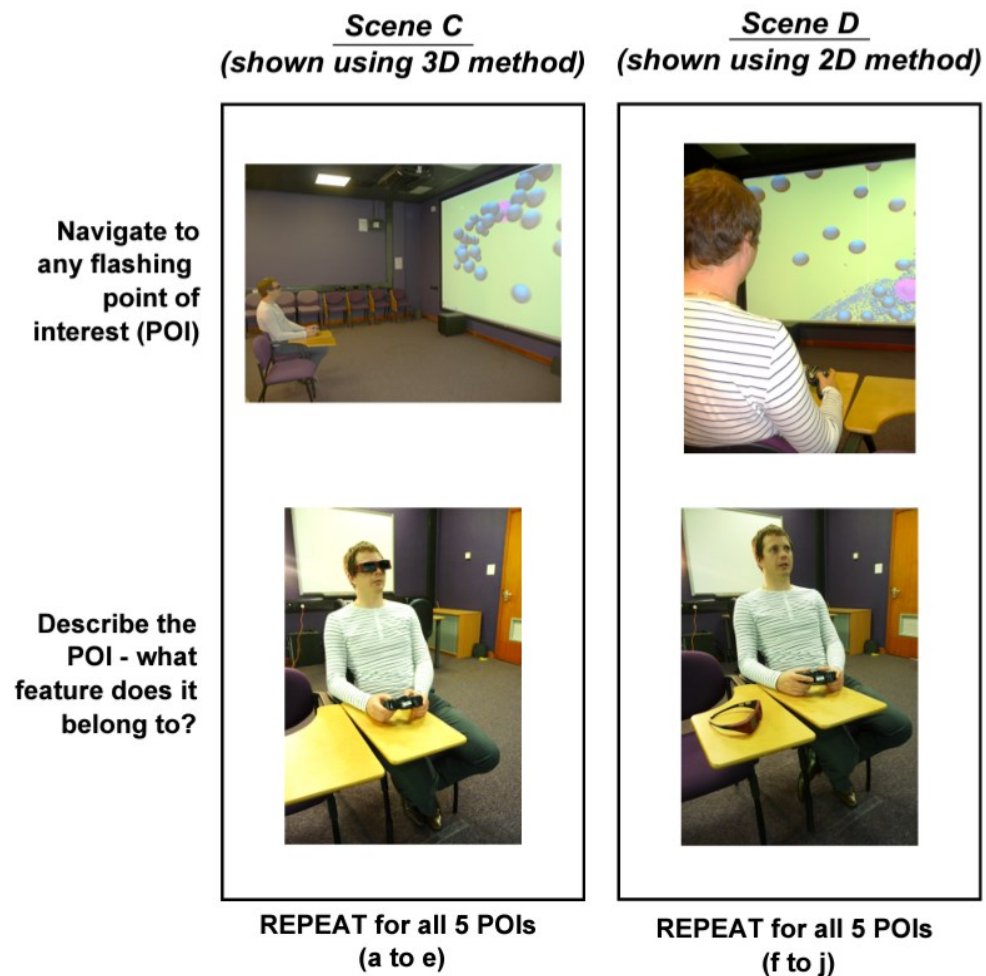


Figure 6-3. Photographs showing virtual reality theatre set-up, during interpretation task. Participant posed by model. Lidar data © Airbus Defence and Space Ltd. (2013).

Scene C : flat, suburban area, for interpretation task

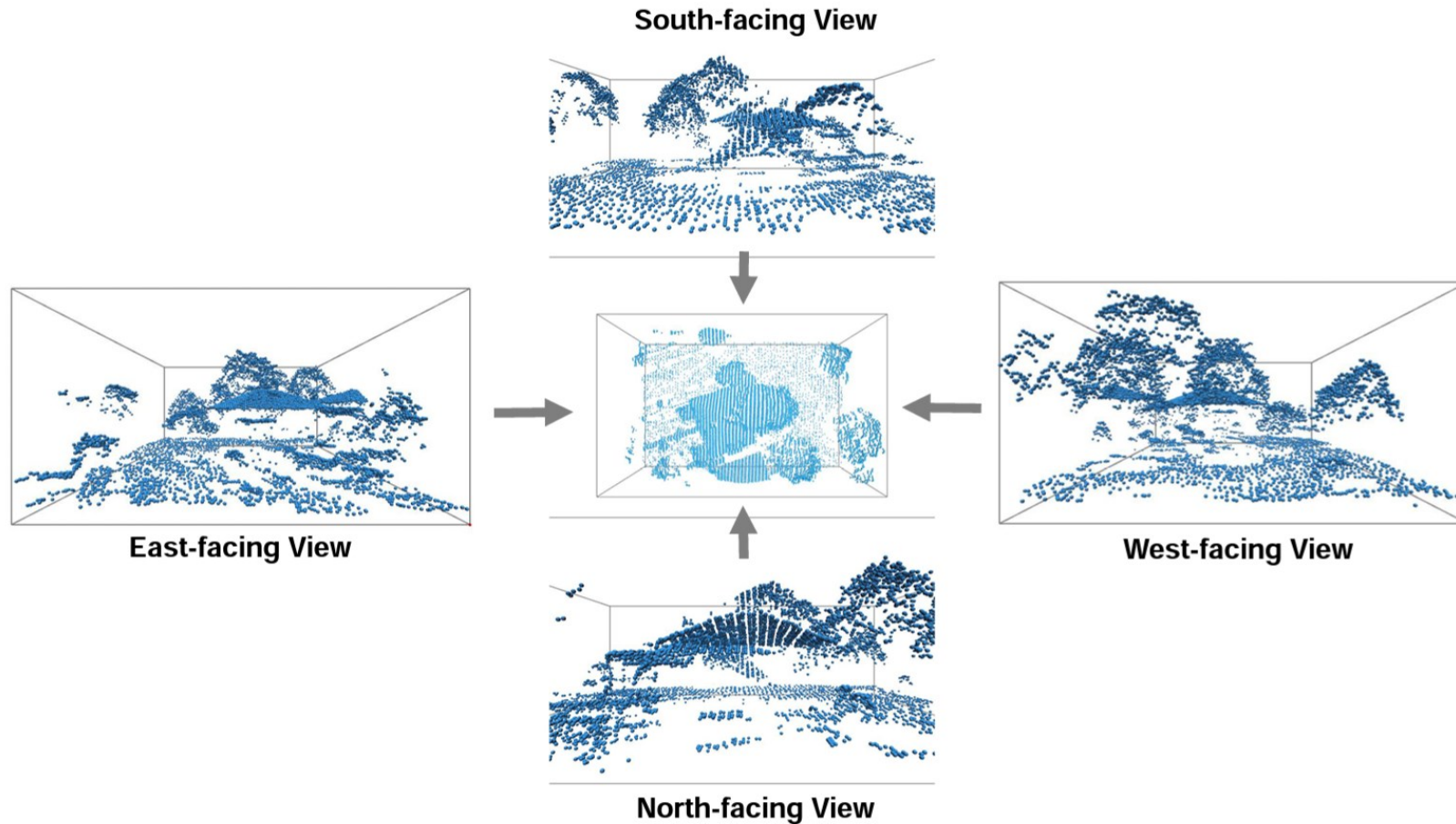


Figure 6-4. Plan and side views of Scene C, which measures (width x length x height) approx. 53m x 35m x 18m. The points cloud is made up of 1563 ground points, 2450 high vegetation points, and 2991 building points. Lidar data © Airbus Defence and Space Ltd. (2013b).

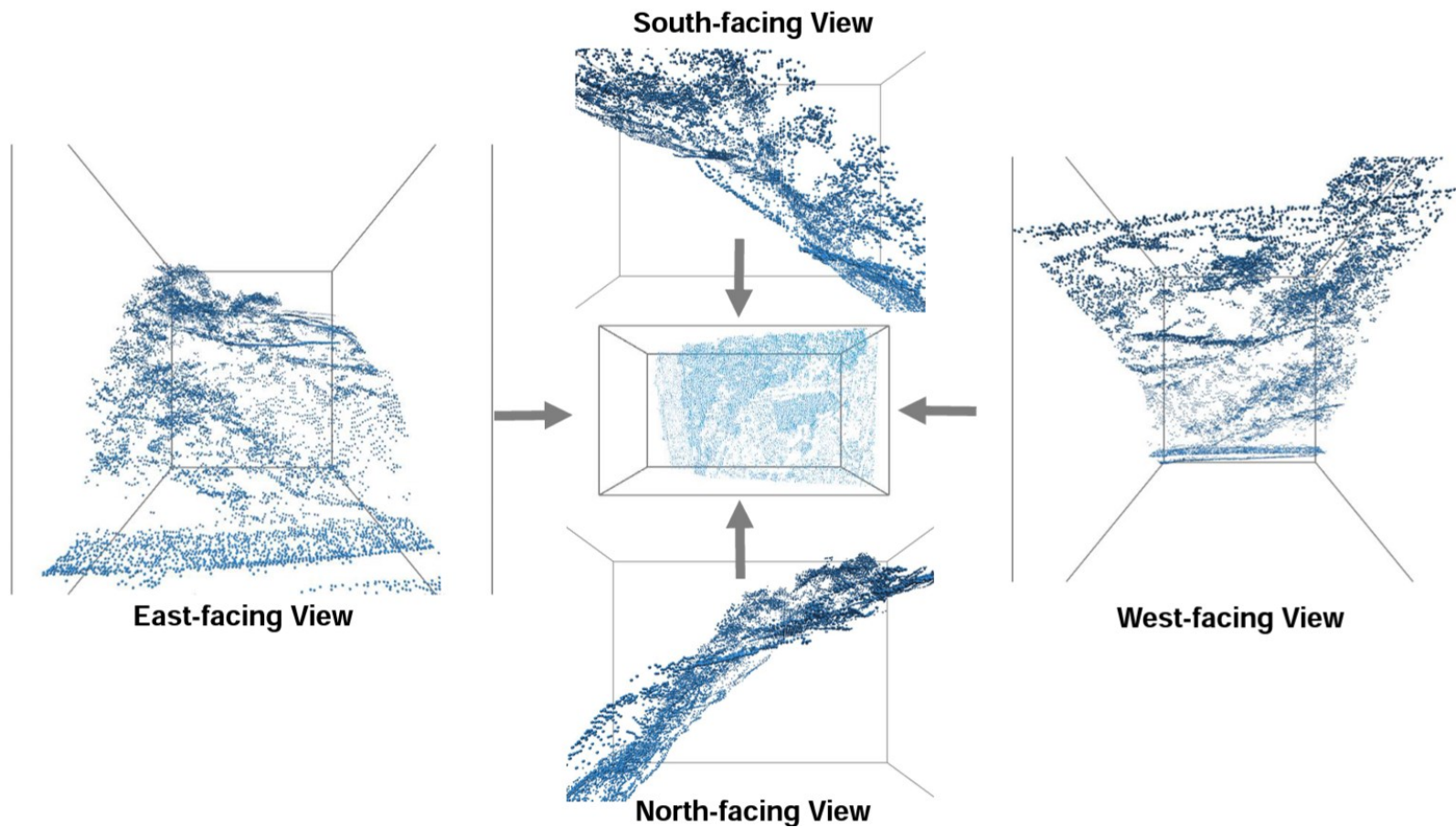
Scene D : sloping valleside, for interpretation task

Figure 6-5. Plan and side views of Scene D, which measures width x length x height) approx. 50m x 35m x 75m. The points cloud is made up of 6209 ground points, 7242 vegetation, and 1123 building points. Lidar data © Airbus Defence and Space Ltd. (2013a).

Qualitative feedback

Also, as part of the feedback section of the experiment, participants were asked Question 10 (of a wider survey), detailed in Table 6-8 which concerned their level of confidence during their interpretation of POIs in each scene. The answers contribute to RQ2.2, regarding acquisition environment complexity. The rankings for the two environment types were collated from both of the scenes to investigate the effect of the visualisations on the participants.

Table 6-8. Question 10, which was posed to participants after completion of the Interpretation task.

| Main Question | Q10 How confident were you about the answers you gave when identifying the feature of the flashing points... |
|------------------|--|
| | in Scene C – flat ? |
| | in Scene D - sloped ? |
| Possible Answers | 1 = Very confident; 7 = Not at all confident, 8 = don't know, 0 [no data inputted] = n/a, |

6.2.3 Data Analysis

Quantifying correctness of classification

The participants' interpretation descriptions were extracted from the transcriptions and checked against notes made during the experiment. These answers were collated in Excel and processed into either a correct or incorrect result for each interpretation, e.g. of a sample of participants who interpreted POI Z (say, a tree) in 2D, 56% was right and 44% wrong. The proportion of correct and incorrect interpretations were ultimately carried through to statistical analysis, which is explained further in the next sub-section. The following text outlines the decisions that were made to convert participants' qualitative responses into a quantitative (in)correct answer.

The volunteers were expected to describe the feature to which the flashing point belonged. However, the participants were not explicitly told to describe the feature to a certain scale or level of detail, so there was a variation in the depth of the answers. Appendix L includes a summary of answers given for each POI in each scene, for each visualisation method (2D or 3D). The POI C (hedge) feature that is described as vegetation at a larger scale are deemed as incorrect. Examples include trellis arch with plants growing, cover to walkway, and an avenue of trees. During their analysis of Scene D, one participant mentioned the covering of a porch or doorway for POI D. This description indicated a man-made feature that is part of the house, so this is deemed correct. Another participant, P35, states “roof” for POI B (chimney), but this is not detailed enough.

With POI F (road), the intention was for the participants to acknowledge there is a man-made road surface within the scene. However, ambiguous terms were used to describe F, including ‘terrace’, whose definition could be natural or man-made, according to Oxford Dictionaries Online (OUP, 2014). Examples of the accepted answers are included in Table 6-9.

For each participant’s interpretation for each POI, a value is assigned according to the accuracy of each answer. A correct answer is given the interpretation score of 1 and incorrect is 0. For POI A, the number of participants who successfully interpreted the feature as vegetation in 2D is simply calculated by summing up the interpretation scores. The resulting number, x , is referred to in the statistical tests in conjunction with the total number of 2D observations, n . This describes the generation of results for the 2D representation of POI A, but the same procedure is carried out for the 3D group for that POI. The generation of x for 2D and 3D is systematically calculated for the two POIs.

The analysis of the resulting quantitative information is described in the section 6.3, which also shows graphical representations of the distributions and statistical relationships between 2D and 3D interpretations.

Table 6-9. Examples of the accepted and rejected interpretation answers for each point of interest (POI), from the Scene C (POIs A-E) and Scene D (POIs F-J).

| POI | Answer | Accepted | Rejected |
|-----|--------------------|--|---|
| A | Vegetation (shrub) | Tree, vegetation, bush, shrubbery, low shrub, | Porch roof, top of bay window, small roof, statue, garden shed, stairs |
| B | Chimney | Chimney, aerial, chimney with moss | Chimney or plant, roof, dormer window |
| C | Hedge | Hedge, wall with plants on, vegetation on wall, rhododendron bush, avenue of trees | Wall, pipes, roof/cover to walkway, man-made raised walkway, trellis arch with plants growing |
| D | Extension | Flat-roof extension, another building adjacent, roof of bits that stick out, porch, roof | Out-house |
| E | Tree | Tree, part of tree, one of two trees, | Roof, edge of roof OR tree, unsure – could be either. |
| F | Road | Road, man-made flat ledge, paved ground, ground flattened by man-made, a flat area, maybe a road or a small field of football field or track or something like that, road or terrace, sea-wall or road | Ground, terrace, ledge, flat part of ground |
| G | Cliff | Cliff face, end of a natural slope , ground surface on steep slope, hill, slope, rock | Waterfall, vegetation, building, man-made |
| H | Vegetation (tree) | Tree, edge of tree, lower part of conifer, christmas-tree-shaped tree | Ledge, on ground beneath tree, ground |
| I | Rooftop | House, roof, building roof | Tree crown, dense vegetation, towards top of rocky outcrop, canopy, part of a tree, boulder |
| J | Rooftop | Eaves of dormer window, roof, man-made structure, where roof dips in (gutter) | Shrub, natural mountainous surface, vegetation, building or mountain |

Test of differences in proportion between two groups

The interpretation data are generated by a sample of the population and some of the observed values will have occurred by chance. Analysis of the data must acknowledge this phenomenon. The participants' interpretations are either correct or incorrect, so the proportion of right answers can be calculated and investigated for bias. During the data analysis stage, this is carried out on the collective results each of the ten POIs for each method (2D and 3D), generating 20 statistical results and plots. The proportion of correct 2D results are then compared to the 3D results for each of the ten POIs.

If the results of the binomial tests and plots suggest data do not follow a Gaussian, bell-shaped curve, non-parametric analysis is required. The 2D and 3D groups are unpaired and not of equal number of observations, so a non-parametric comparison of the two was made through a chi-square test. This is applied to a 2x2 contingency table of results in the R package (R Core Team, 2014), with n values for 2D and 3D in one row and their p values in the other, as shown in Equation 3.

Equation 3.

```
prop.test(x = c(smallgroup[i,x],biggroup[i,x]),
n = c(smallgroup[i,n],biggroup[i,n]),
alternative,
conf.level = 0.95)
```

Where:

x is the number of successes,

n is the number of observations,

p is the probability of the success (out of a maximum of 1),

and `alternative` refers to the user-specified type of test, where the first group is less than, not equal to (two-tailed), or greater than the probability of success in the second group.

Within the function, the variables of the group of the lowest n , (`smallgroup`), must be followed by those from the group with the higher n (`biggroup`). Therefore,

Scene C was set at 2D (smallgroup) less than 3D (biggroup) and used H₂. Scene D was set at 3D (smallgroup) greater than 2D (biggroup) and followed H₃. Ultimately, the two types of analyses pose the same question – is 3D better than 2D? The results shown are not adjusted according to Yate’s continuity correction (Yates, 1934).

The hypotheses tested are explained below:

Test of Equal Proportions

Null Hypothesis H₀: $p = 0.5$

No bias towards the probability of one group being more or less successful than the other group, *i.e.* there is a ½ (fair) chance that probability of success in the first group (say, 2D) is the same as the (3D) second group.

“Two-tailed” Alternative Hypothesis H₁: $p \neq 0.5$

There is bias towards the probability of one group being more or less successful than the other group, *i.e.* there is not a ½ (fair) chance that probability of success in the first group (say, 2D) is the same as the (3D) second group.

“Less than” Alternative Hypothesis H₂: $p < 0.5$

The probability of the smaller group being successful is less than a ½ fair chance, *i.e.* there is not a ½ that probability of success in the smaller group is less than that of the larger group.

[applied to Scene D]

“Greater than” Alternative Hypothesis H₃: $p > 0.5$

The probability of the group being successful is greater than a ½ fair chance (positively skewed), *i.e.* there is not a ½ fair chance that probability of success in the smaller group is greater than that of the larger group.

[applied to Scene C]

Analysis of bias

A binomial test is used to check for randomness in the data (Dytham, 2006, p. 220) and whether its distribution fits a normal Gaussian frequency curve. The test assesses the proportion of correct interpretations for n number of participants (Myers *et al.*, 2010) and its underlying assumptions are outlined by Myers *et al.* (2010, p. 70), as follows:

- a) Only two outcomes / a two-class population.
- b) The proportion of correct interpretations is the same.
- c) The probability of each trial outcome does not rely on any other trial.

The experimental data met these criteria – (a) the interpretation outcome was either correct or incorrect, (b) the null hypothesis expected these proportions to be equal, and (c) the results did not rely on any other trials. The `binom.test()` function in R (R Core Team, 2014), Equation 4, was therefore used to determine the likelihood of achieving a null hypothesis, in which the outcome of success and failure are the same.

Equation 4.

```
binom.test(x, n, p = 0.5, alternative, conf.level = 0.95)
```

Where:

x is the number of successes,

n is the number of observations,

p is the probability of the success (out of a maximum of 1),

and `alternative` refers to the user-specified type of test, where the first group can be less than, not equal to (two-tailed), or greater than the probability of success in the second group.

The 20 observed p-values for each group's POI interpretation were independently assessed in R using this technique. If a p-value is above 0.05 (95% confidence level), the null hypothesis is accepted, indicating is a fair

chance of participants being correct. An alternative hypothesis is accepted when the p-value is less than 0.05 and the chance of being right or wrong is not equal. Here, H_1 assumes the results are skewed towards being correct. A description of the null and alternative hypotheses tested follow:

Exact Binomial Test

Null Hypothesis $H_0: p = 1/2$ or 0.5

The group is not biased towards being correct *i.e.* there is a $1/2$ (fair) chance of success.

“Greater than” Alternative Hypothesis $H_1: p > 0.5$

The probability of the group being successful is greater than a $1/2$ (fair) chance and has a positively skewed distribution curve, *i.e.* there is not a $1/2$ (fair) chance of success.

Quantitative and qualitative audio analysis

Similarly to the measurement task, themes were identified within the transcriptions by allocating selected text to nodes. Specific comments are included in Scene C and D results. The responses to the Feedback question (Q10), regarding the answers giving when identifying the flashing points, are displayed graphically to help understand participant confidence for the interpretation of each scene.

6.3 Results

The number of participants that took part in the interpretation task is detailed in Table 6-10 and participant background information can be referred to in Chapter 4. A table summarising interpretation accuracy results generated by these participants that were generated from the quantification of verbal answers

are shown in Appendix L, alongside examples of 2D and 3D answers for three POIs. In this section, for each scene, a description of the measurement results for the 2D and 3D groups are presented, alongside normality and significance testing. Qualitative themes that were mentioned by the participants both during and after the trial are presented in the results. A between-method (2D vs. 3D) comparison is made to understand the differences, if any, between the methods for each POI. The likelihood of a correct interpretation for each of the POI features is revealed by considering binomial distribution plots and tests for 2D and 3D participants.

Table 6-10. Number of participants taking part in the interpretation task.

| Scene | AOI description | No. of participants | |
|-------|----------------------------|---------------------|-----------|
| | | 2D method | 3D method |
| C | Suburban area (London, UK) | 22 | 24 |
| D | Valleyside (Bristol, UK) | 24 | 22 |

6.3.1 2D vs. 3D interpretation of features (POIs A – J)

The results for the comparisons between 2D and 3D groups for each POI (A-J) are shown in Table 6-11, which describes a two-tailed test of the difference in proportions. The null hypothesis states that the proportion of correct interpretations for 2D and 3D are equal; a probability result (p-value) of <0.05 agrees with this null hypothesis. In Table 6-11, the highlighted POIs (shrub (A), hedgerow (C), and building (I and J)) have a p-value of 1, which tells us that at a 95% confidence level, there is 100% probability that the proportions of success in 2D and 3D groups are the same. We can therefore assume that there is a fair chance of achieving the same results in 2D and 3D for these particular features. The possible reasons for this are discussed in section 6.4.

The research also determines whether the POIs have more accurate interpretations in 3D, in comparison to 2D (although we have already

established $2D = 3D$ for A,C,I,J). shows the probability of success in the 3D groups are greater than that of the 2D group, at a 95% confidence level. These results are key to the answering RQ2.1b. The accuracy of interpretation of POIs B (chimney) and G (cliff-face) are better when using the 3D display method. If we lower the confidence level to 90%, two more features show answers that are proportionally more correct in the 3D method, compared to the 2D method. These are E and H, which are both vegetation features, from Scene C and D respectively.

The description column in Table 6-11 provides descriptive accompaniment to Figure 6-6, which graphically summarises the differences between the 2D and 3D methods for the individual POIs. These are the key results of the Interpretation task.

Comment on trends

The plots in Figure 6-6 suggest that the road feature in Scene D (POI F) displays a 12% higher interpretation accuracy in the 2D group. However, when studying Table 6-11, the statistical analysis for $2D > 3D$ confirms that none of the POIs has a higher 2D interpretation. There are several features that have similar results; for POI I and J, in both methods, the participants struggled to understand the feature in 2D and 3D (all ranging between 42% - 50%). The hedge feature of Scene C (POI C) had a low result (<40%) for both methods. Equally, similar higher accuracy results were seen for POI A (shrub) and D (extension). The POIs with a significant difference in interpretation accuracy were B (chimney), G (cliff), and two trees (POI E and H). This shows that, under the experiment conditions, 3D visualisation generates better interpretation results than 2D for the same task. However, this advantage may not be required if the end-user is satisfied with the adequacy of the 2D degree of accuracy. This matter is revisited in section 6.4.1.

Table 6-11. p-value results of 2D vs. 3D interpretations for each POI (A-J), using a two-tailed 2-sample test for equality of proportions with continuity correction (prop.test() function in R (R Core Team, 2014)). Bold p-values denote a significant result between the 2D and 3D groups, at 95% confidence level, i.e. <0.05.

| p-values for <u>probability of success</u> for 2D vs. 3D groups under conditions below: | | | | | | Outcome of 2D vs. 3D interpretation results | |
|---|-----|-----------|------------|--------|------------|---|-------------------|
| Scene | POI | Feature | 2D greater | Equal | 3D greater | Description | Summary |
| C | A | shrub | 0.500 | 1 | 0.500 | The shrub located at the front of the house in Scene C had similar results (2D = 3D, 73%:75%). | Same |
| | B | chimney | 0.974 | 0.0518 | 0.026 | There was a significant difference in 2D and 3D interpretations of the chimney feature (2D < 3D, 64%: 92%). | 3D better at 0.05 |
| | C | hedge | 0.500 | 1 | 0.500 | The garden hedge returned the worse interpretation results overall and the accuracy of 2D and 3D is the same (2D = 3D, 36%: 38%). | Same |
| | D | extension | 0.518 | 0.964 | 0.482 | The two-tailed test revealed a high probability, 0.96, that there is no difference between 2D and 3D results for the house extension (2D vs. 3D, 95%: 100%). | No sig. outcome |
| | E | tree | 0.927 | 0.145 | 0.073 | This POI, in Scene C, is located on a very ambiguous interface between the corner of a house and a nearby tree. The 3D answers were more often correct than the 2D interpretations, but not significantly so, at 95% level (2D < 3D at 90%, 41%: 67%). | 3D better at 0.1 |
| D | F | road | 0.300 | 0.599 | 0.700 | This POI was the only feature that showed a better percentage of interpretations in the 2D group, compared to the 3D group. The man-made flat, road feature was more obvious in 2D than 3D, although this is not proven significantly (2D vs. 3D, 71%: 59%). | No sig. outcome |
| | G | cliff | 0.960 | 0.070 | 0.035 | The 3D interpretation of the cliff-face yielded more accurate results than 2D answers (2D < 3D, 71%: 95%). | 3D better at 0.05 |
| | H | tree | 0.936 | 0.129 | 0.064 | A higher proportion of 3D participants were able to correctly classify the tree , than the 2D group (2D < 3D at 90%, 75%: 95%). | 3D better at 0.1 |
| | I | roof | 0.500 | 1 | 0.500 | Many participants recognised that POIs I and J were part of the same feature, but the 2D and 3D results were not significantly different (2D = 3D, 46%:50% for I and 42%:52% for J). | Same |
| | J | roof | 0.500 | 1 | 0.500 | | Same |

Proportion of Correct Feature Interpretation for 2D and 3D Participants Groups

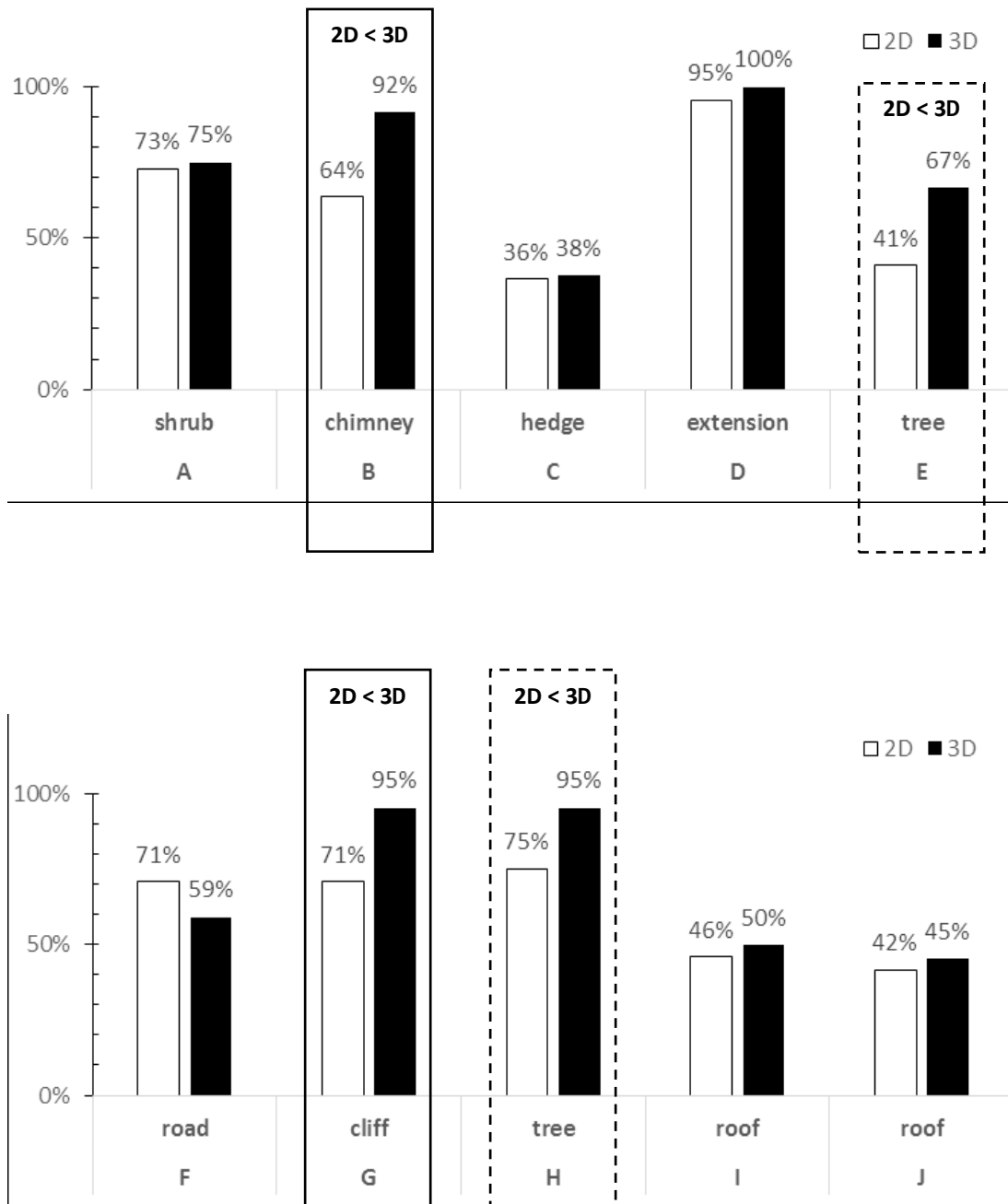


Figure 6-6. Plots comparing the interpretation results (accuracy out of 100%, y axes) for 2D vs. 3D for each POI (x axes). Solid rectangle denotes significant difference in proportionality between 2D and 3D interpretation results. Dashed rectangle describes the same, but at a 90% confidence level.

Analysis of bias

While analysing the participants' interpretation accuracies, it is important to understand whether there was any underlying bias in the interpretation of the POI. Binomial distributions describe whether the interpretation answers (correct or incorrect) are biased towards being correct. This analysis can only be examined as a product of the accuracy results. The null hypothesis states there is no bias between success and failure; the frequency distribution would be close to a normal, bell-shaped curve. Here, since frequency probability is a function of n and the number of participants is uneven among the 2D/3D groups, the binomial distribution results are broken down into the 2D and 3D methods for each scene. The interpretation of Scene C using the 2D method are summarised in Table 6-12. The p-values represent the chance (out of 1 or 100%) of a null hypothesis (fair outcome), so if the p-values are less than 5%, one can be confident that the proportion of incorrect:correct interpretations for those particular POIs is unlikely to be attributed to chance.

Firstly, for Scene C, Table 6-12 shows that only POI B exhibits bias towards a correct interpretation in the 3D method and not the 2D method, which is unsurprising for techniques that are deemed significantly different. Regarding Scene D, Table 6-12 shows that both 2D and 3D methods displayed low probabilities that their proportion of correct answers occurred by chance. This suggests that the road feature (F), the cliff-face (G), and the tree (H) were biased towards being accurately interpreted by the participants from the 2D and 3D groups. G and H had particularly low p-values (5.48E-06 for both) in 3D, which infer a higher frequency of positive results in the 3D group for these POIs.

Regarding interpretations that returned a higher proportion of incorrect results, only POIs I and J (situated on the building at the top of the hillside) were not biased towards being correctly interpreted in 2D or 3D. The interpretation results of these features, in both methods, could have occurred by chance (p-values for 2D: 0.458 (I) and 0.417 (J); 3D: 0.500 (I) and 0.455 (J)). POI G, however, displays bias in both 2D and 3D interpretations. Despite both being significantly biased, their proportions are different enough to be significant.

Table 6-12. Binomial distribution results for 2D and 3D interpretation of POIs A to J. Where x is the number of correct interpretations, n is the total number of observations, and p is the probability of success (x/p). P -value is given for the results of the “success is greater”-tailed `binom.test()` performed in R, at the 95% confidence level. POIs with p -values below 0.05 display bias towards being correct and are shown in bold. Normal distribution for the given n is also shown.

| Scene | POI | Feature | p-value | |
|-------|-----|-----------|----------|----------|
| | | | 2D | 3D |
| C | A | shrub | 0.026 | 0.011 |
| | B | chimney | 0.143 | 1.79E-05 |
| | C | hedge | 0.933 | 0.924 |
| | D | extension | 5.48E-06 | 5.96E-08 |
| | E | tree | 0.857 | 0.076 |
| D | F | road | 0.032 | 0.262 |
| | G | cliff | 0.032 | 5.48E-06 |
| | H | tree | 0.011 | 5.48E-06 |
| | I | roof | 0.729 | 0.584 |
| | J | roof | 0.846 | 0.738 |

In the cases where POIs are biased in both 2D and 3D (A, D, F, G, H, which make up 50% of the Interpretation experiment), the unexpected high accuracy of human interpretation could be an effect of a confounding factor, which has overridden any effects of the display techniques. This is an important factor to take into account and it is revisited in section 6.4.

Trends in qualitative feedback

Comments were considered as auxiliary information, to help explain the quantitative results for RQ2.1. It became clear that the participants required an understanding context of wider environment ahead of determining the feature classification. The following excerpts illustrate how some of the volunteers used

their knowledge of the two environments to assist their interpretations of the features from the point clouds. This contextualisation technique is used in both 2D and 3D.

“[POI A] It's in about the right place for a chimney and trees don't grow out of the top of houses. It doesn't look like an extension of the big tree, I suppose it could be, but you tend not to have trees over... So just going on, kind of, cultural/environmental memory really.”

Scene C in 3D, P23

“You can't really tell the shape from it, and because I think the rest is like a house and stuff, then there's probably going to be trees around.”

Scene C in 2D, P39

“On this side, it looks like this is on the side of a mountain and there are trees. Looking at it from the top, the middle now has... it might look like a park, a national park, and the trees are on one side...”

Scene D in 2D, P36

“Oh! That one [POI G], that's the part that I said it looked like a capped part of the mountain, because it's very, very steep from the top of the view, so that's the part I said that probably has been cut off for building of the road.”

Scene D in 3D, P41

During interpretation, the participants applied the knowledge gained from the lidar summary that was taught prior to the experiment. Where aspects of this information conflicted, *i.e.* points I and J, the participants had to use their own judgement.

“This looks like a small tree [...] because of the shape and that there’s ones below it, so like it’s going through.”

POI H, Scene D in 3D, P44

The volunteers also referred back to Scenes A and B, which were used during the measurement task, and used these as a reference guide. Scene C was similar to Scenes A and B.

“It could be a mini-roof, you know like we saw earlier, coming off the edge of a building”

POI A, Scene C in 2D, P44

“There seems to be a... similar to the building that was the second point cloud which had a lower roof lines”

Scene C in 2D, P09

6.3.2 Effect of physical AOI (Scenes C & D)

The effect of the physical complexity of each AOI (and its subsequent representation) on participants’ feature interpretation was explored using quantitative and qualitative feedback. The quantitative results are the mean proportion of correctly 2D and 3D POI answers for each scene, which are shown in Table 6-13. For each scene, it appears that the 2D results (C: 62% mean POI accuracy, D: 61%) are lower than their 3D equivalent interpretations (C: 74%, D: 69%), however, ultimately, 2D vs. 3D comparison of the mean AOI accuracies was not considered as part of the results because of POI selection technique. This is further discussed later in the chapter, during reflection of the methodological development.

Table 6-13. Summary of mean interpretation accuracy per AOI, based on mean POI accuracy

| | Mean interpretation accuracy result | | |
|------------|-------------------------------------|-------|-----------|
| | 2D | 3D | 2D and 3D |
| C – flat | 0.618 | 0.742 | 0.68 |
| D - sloped | 0.608 | 0.691 | 0.64 |

As part of the Feedback Section of the experiment, participants ranked their level of confidence during their identification of POIs and the different types of features. Regardless of technique used, overall, the participants felt particularly confident in identifying the flashing points in Scene C, as shown by the skewness towards 1 (very confident about flashing point answers) in Figure 6-7. In comparison, Scene D has a more normal distribution, with answers ranging range from very confident to not at all confident. Although these results could be further split into 2D and 3D for each scene, to help answer RQ2.2, the further breakdown of results were potentially biased by the unpaired testing of 2D and 3D tasks. For this reason, the 2D vs. 3D feedback results are not considered for RQ2.2.

Q10 How confident were you about the answers you gave when identifying the feature of the flashing points for Scene C and D.

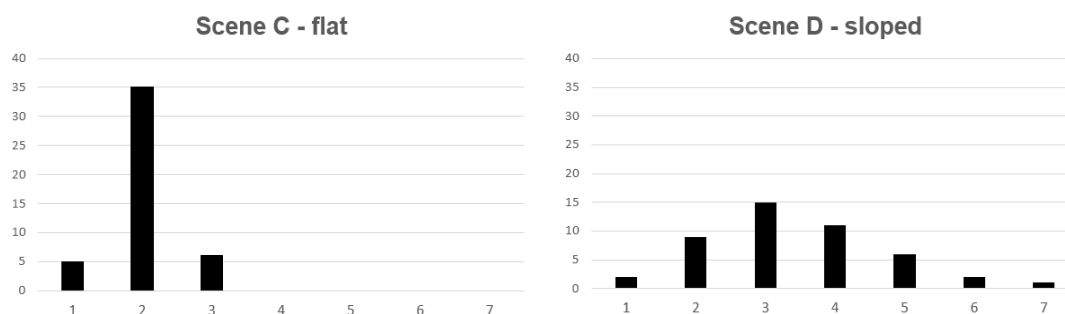


Figure 6-7. Graphs showing the overall level of participant confidence while carrying out the interpretation tasks for Scene C (left) and Scene D (right). Scores were given in response to Feedback Q10. Results represent the general combined confidence scores for both 2D and 3D methods, for each Scene. X axis scale: 1 = very confident..., 7 = not at all confident about the answers given when identifying the feature of the flashing points. Y axis: % of participants (total of 46 for each scene).

Participants commented further on the scores that they provided for Q10 and quotes regarding Scene C can be found in Appendix L. According to the extracted quotes below, Scene C was similar to Scenes A and B, and a more familiar environment to the participants.

“Okay, the environment of C was... I was more confident in, so maybe a... 2. I think maybe it was just more similar to the other ones we’d already looked at.”

C in 2D > D in 3D, P40

“I was more confident in the flat scene than the sloped scene, so in the sloped scene. I think I was still confident in identifying the points, but... definitely less so because it was sloped. And I think that’s because I don’t see hills very often <laughter> so I wouldn’t be necessarily [able to] pick out features because I don’t see them everyday. But, like I said, I see trees everyday, and buildings, so they’re a bit easier to identify. [...] Scene D is clearly... it’s quite like a steep slope, so... I don’t know features that are common on steep slopes to be able to identify them in point cloud form.”

C in 2D > D in 3D, P21

When considering the individual features within each environment, in the flat Scene C environment, the man-made POI D (building) was extremely easy for both 2D and 3D groups to interpret (all achieving 95-100% correctness). However, in the sloped Scene D, the rooftop on which I and J were located was unexpectedly irregular, compared to the buildings in Scene C and A. The sloped nature of the environment may be less of a factor in the interpretation of POIs I and J, instead the uncharacteristic lidar returns could be the main reason for the low proportion of correct answers (42% - 50% across 2D and 3D interpretations of I and J). There were considerably more remarks about the sloped scene, Appendix L– ‘Further Comments for Q10’, predominantly

regarding POIs I and J. The contrasting characteristics of the points – jumbled/irregular structure and the lack of points underneath the feature, made the two POIs difficult to understand. The elevation of the scene also caused issues (see Appendix L) and made it confusing for participants.

“Maybe on a different scene, I would be MORE confident. If it was similar to the first house we looked at [in Scene A], if this last scene looked like that, then I would be... 100% confident it’s a house, whereas this one... it’s more uncertainty in the actual scene.”

Scene D in 2D, P28

“I think I was slightly less confident with that one, I think, in getting features confused between what was actually ground on the slope or raised foliage on the slope.”

Scene D in 2D, P34

6.3.3 Results summary

The interpretation task presented 10 features of varying degrees of difficulty, represented by one flashing point each, and 2/10 of the manual classification of features were significantly more accurate in 3D, see Table 6-14. The remaining eight features had similar accuracies in 2D and 3D, whether high or low.

Participants felt that they had a better understanding of the five (of the ten) POIs in Scene C, the suburban area, because it was closer to their daily, real-life environment. Although 2D vs. 3D comparisons were not made between the AOIs, members of the 2D group stated that the incline of Scene D made its interpretation confusing. Additional findings, relating to both 2D and 3D methods, are listed below Table 6-14.

Table 6-14. Summary of the relationship between 2D vs. 3D interpretation accuracy for ten POI features and participant feedback regarding each scene. There were no cases where 2D > 3D.

| Scene | POI | Feature | Summary of quantitative statistical comparison | | | Summary of participant feedback from 2D & 3D groups |
|-------|-----|-----------|--|---------|---------|--|
| | | | 2D > 3D | 2D ≈ 3D | 3D > 2D | |
| C | A | shrub | | ● | | - Considered a familiar, everyday scene. - Confident that interpretations for POIs are accurate. |
| | B | chimney | | | ● | |
| | C | hedge | | ● | | |
| | D | extension | | ● | | |
| | E | tree | | ● | | |
| D | F | road | | ● | | - Unfamiliar with features that might be on slope. - Low confidence in accurate interpretations for POIs. - 2D participants commented: confusion between ground and raised vegetation on ground. |
| | G | cliff | | | ● | |
| | H | tree | | ● | | |
| | I | roof | | ● | | |
| | J | roof | | ● | | |

- Participants felt it was essential to understand the context of the environment, before deciphering the highlighted points from within the larger cloud.
- Participants applied knowledge learnt from the lidar overview. Similarities were acknowledged between the features within Scene C and previous examples (Scenes A and B), but Scene D was unfamiliar.
- Accuracies could not be compared between Scenes C and D because of POI selection technique. Unbiased direct comparisons of 2D vs. 3D participant rankings for Scenes C and D were not possible because of the unpaired sampling.

- The verbal classification technique meant that some participants provided more detailed answers than others, but when translation was required, the interviewee's intended meaning was robustly verified, as a result.

6.4 Discussion

Research Question 2 asks whether visual interpretation of lidar point cloud features (with varying geometries) is more accurate when carried out via 2D or 3D visualisation. Comparisons were made between 2D and 3D visualisation of points derived from different types of features, which were located in two different environments. The 2D vs. 3D results are described in relation to the sub-research questions.

6.4.1 2D vs. 3D interpretation of different features (RQ2.1)

The accuracy of the point cloud interpretation was calculated for RQ1.1a, as a pre-requisite to RQ2.1b. The accuracies of the POI interpretations are discussed between the 2D and 3D groups.

RQ2.1 Is lidar point cloud feature interpretation more accurate when carried out via 2D (monocular) or 3D (biocular) visualisation?

(a) what is the 2D and 3D interpretation accuracy?

(b) is there a significant difference between interpretations of the point cloud features made in 2D, in comparison to those made in 3D?

RQ2.1b) Differences between 2D vs. 3D feature interpretation accuracy

RQ2.1(b) requires the research to establish whether the accuracy of 3D feature interpretations are better than 2D. An objective of the research was to show the volunteers POIs derived from a range of point cloud feature geometries, to help investigate which kinds of features might benefit from which method. Therefore

it is expected that, across the POIs, there will be a variation in disparity between the accuracy of 2D and 3D interpretations. The schema in Figure 6-8 summarises the results and the following discussions related to the POIs that had similar 2D and 3D interpretation accuracies, statistically equal accuracy proportions, and any whose feature identification were better in 2D, as opposed to 3D.

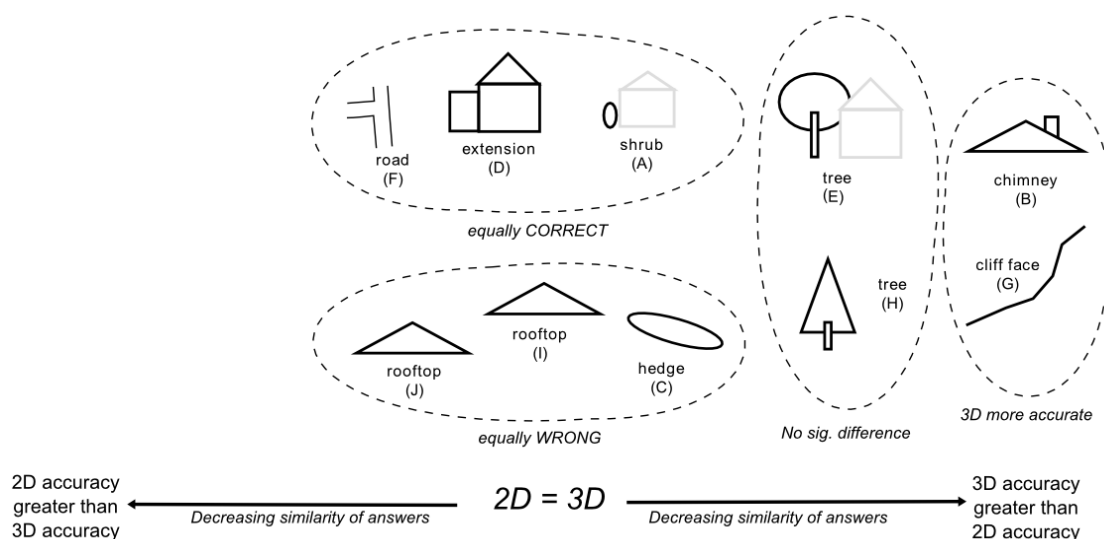


Figure 6-8. Interpretation Task findings, showing the difference between 2D and 3D interpretations of the ten points of interest. There were no features that had a better 2D accuracy than 3D.

Features with Similar 2D and 3D Accuracies

A shrub (A), hedgerow (C), and building (I and J) were all 100% likely to yield the same results in 2D and 3D, whether correct or incorrect. In Figure 6-8, these features shown are clustered within the 2D = 3D central region, but are separated into those that are equally correct and equally wrong.

Equally Correct. It was anticipated that the shrub (A) would be difficult to understand because it was in close proximity to what most participants identified as a house and it was of a quasi-man-made/natural structure. However, results from the 2D group and the 2D group were similar (2D:73%, 3D:75%) and showed the chi-squared-based comparison of proportions shows no variation at 95% confidence level. It was predicted that participants would benefit from the spatial understanding of lidar provided by the depth of the 3D

method, as claimed by Kreylos *et al.* (2006). The reason the other points were interpreted similarly may be because the monoscopic clues (Hoffman *et al.*, 2008) provided sufficient depth perception (Lai *et al.*, 2010; Neves *et al.*, 1997), e.g. the shadow on the points provided enough of an impression of three-dimensionality. Two planar features are also correctly interpreted by both groups F and D. The extension in the suburban scene (POI D) and the road feature in the inclined environment (POI F) had differences between 2D and 3D interpretation accuracy, but they were not strong enough to be significant. Interestingly, both features were man-made, flat surfaces.

Equally Wrong. The features that returned equally wrong answers from both the 2D and 3D groups are C (hedgerow in Scene C) and POIs I and J (part of same rooftop in Scene D). The results for hedgerow (C) were statistically alike (2D:36%, 3D:38%) and their low percentages indicate it was difficult to interpret in both methods. However, the participants did not have a selection of potential classifications, as is compulsory in citizen science, when multiple users are analysing data to the same standard (Foody *et al.*, 2015). Neighbouring points were arranged in a linear fashion, seemingly projecting from the house in Scene C, and, without classification guidance, many participants suggested the feature was a pipeline or corridor/walkway structure. The I and J points were located on different parts of a rooftop (Scene D), whose lidar returns were uncharacteristically messy and in close proximity to similarly irregular vegetation points. It is assumed that that participants, of mixed lidar background, were unable to recognise building in scene.

When reflecting on the similar 2D and 3D results, it becomes apparent that the following characteristics may have an overriding effect on human interpretation of lidar point cloud features:

- Feature is easy/difficult to recognise.
- Monoscopic cues sufficient in 2D.
- No classification categories available.
- Anomalous return characteristics, e.g. irregular lidar returns from smooth feature.

Features without significant difference in accuracies

POI E was purposefully chosen to coincide with the corner of a building and its neighbouring tree. Its location is highly ambiguous, as acknowledged by the lidar expert who helped to visually validate the POIs, but most participants used the edge of the roofline as a reference surface to determine the POI feature. It is thought that the linear edge adds linear perspective (Howard & Rogers, 2012) to the arrangement of points, which enhanced pictorial cues necessary for understanding of the geographical data (Kraak, 1998). It was anticipated that the 3D results would be more accurate than 2D results, but, at a 95% confidence level, the 3D answers were no better than 2D. POI H is situated on the base of a typical coniferous, conically-shaped volume of points (Scene D), uninhibited by any other features, and the points stereotypically penetrate through to the ground below it. It was expected that the point cloud representation would be easy in both methods because it formed part of such a well-defined shape, facilitating visual interpretation sub-tasks of identification and judgment (Bianchetti & MacEachren, 2015) in both 2D and 3D. The results of the interpretations show that there was high proportion of correct answers in 2D (75%) and higher in 3D (95%), although the 3D result was not significantly greater than the 2D score. The form of the feature may not warrant the addition of stereo depth provided by 3D visualisation.

- Comparison with neighbouring planar features can assist depth cues.
- Well-defined point cloud representation may not warrant 3D depth.

Features with better 3D accuracy

The null hypothesis states that 2D and 3D interpretation results are not significantly different (at 95% confidence level). Of all ten POIs, an alternative hypothesis is accepted for only two features - the chimney feature in the suburban AOI (POI B, Scene C) and the cliff-face in the sloped AOI (POI G, Scene D). The difference between 2D and 3D interpretation accuracy of the chimney feature (2D < 3D, 64%: 92%) was unexpected because it was assumed that both participant groups (2D and 3D) would find identification easy

because of its location, on top of a roof. However, POI B was in close proximity to vegetation and the chimney not exhibit a regular 'chimney shape'. The latter characteristic caused uncertainty in both methods, but the 3D group were ultimately more accurate in their answers, despite the unexpected lidar returns. The intricacy of the spatial arrangement of the points may have demanded the viewer to inspect the feature more closely, increasing the stereo depth cue (Piryankova *et al.*, 2013). POI G is located on a cliff-face and those who explored the point cloud in 3D were more often correct than the 2D group (2D:71%, 3D:95%). The POI is located on a rough, near-vertical plane so participants may have used this contextual information to identify a meaningful interpretation. The surface of the cliff face may not have provided a sufficiently smooth enough surface to enable the 2D participants to perceive a continuous surface of points, which helps determine dimensionality (Wilcox & Duke, 2005).

This study randomly allocated participants with good stereoacuity into two different experiment groups, see Figure 4-11. The 3D participants' interpretation ability was more enhanced than those in the 2D group. The types of features that benefitted from stereoscopic 3D visualisation are likely to be effect on human interpretation of lidar point cloud features:

- Complexity of point arrangement benefits from stereoscopic projection.
- Participants' near-view inspection.

Although several POIs were better interpreted in 2D, one must consider on a case-by-case basis whether the better accuracy offered by 3D is warranted. Depending on the requirements of the end-product (e.g. DEM generation at a certain resolution), a higher accuracy of an immersive stereo 3D visualisation may not be necessary.

6.4.2 2D vs. 3D interpretation of different environments (RQ 2.2)

The 2D vs. 3D results of the experiment were not sufficient to answer the proposed RQ2.2 research question, repeated overleaf.

RQ2.2 Does the complexity of the lidar point cloud acquisition environment affect 2D vs. 3D results?

This was caused by bias introduced by the unpaired testing, which was chosen (over paired testing) to reduce the bias of the learning effect. The group who carried out the 2D results of Scene D also perform Scene C in 2D. The other group did D in 3D and C in 2D. This meant that Scene D 3D was ranked by participants in comparison to the other scene they had viewed, *i.e.* Scene C in 2D, and *vice versa*, creating a skewed, incomparable result. However, the qualitative data highlighted observations in 2D and 3D for each scene.

In Scene C, POIs A-E, were all located on a flat, suburban environment and it was expected (H2.2a) that interpretation would be straightforward for both 2D and 3D participants. Results show that participants felt confident about their interpretation answers within the scene because it had familiar, everyday features. This can be likened to the interpretation sub-task of ‘comparison’, which involves the viewer comparing the presented remotely-sensed data to reference information (Bianchetti & MacEachren, 2015). In addition to the participant observations of recognisable features, the flat terrain meant they were well-defined and relatively segregated.

In contrast to the everyday environment of the flat scene, the slope scene presented an additional challenge of features on an incline and it was anticipated (H2.2b) that the task would be difficult in 2D, but less so in 3D. Participants stated that the elevation caused confusion in differentiating between features situated on the ground surface and the ground surface itself. However, a planar feature within a complex environment (POI F) proved to be easy to understand in both 2D and 3D. Again, this could be attributed to the visual perception of a surface that is interpreted from cognitive interpolation between the continuous points (Wilcox & Duke, 2005). The expected irregular nature of another man-made feature in Scene D is thought to have caused the similar 2D and 3D accuracies for POIs I and J, but the scene’s incline may have exacerbated this effect in both methods, as it would under automated processing (Meng *et al.*, 2010).

6.5 Reflection on Methodological Development (RQ3)

6.5.1 Selection of POIs

Alongside the general accuracy results, the amount of bias towards the results being correct was explored. POIs with a range of difficulties were chosen, so expect accuracy to be similar in the easy- and hard-to-interpret POIs. For half of the POIs, participant interpretation for both 2D and 3D had a distributional skewness towards being correct. This implies that, regardless of method, there is an underlying reason for this skewness, including (a) the POIs are easy/hard to identify, (b) the participants are skilled at point cloud interpretation, or both.

The first task complemented the second by using two contrasting features that were present in both of the interpretation scenes. The participants commented that the features in Scene C were similar to those in Scenes A and B. This was expected because Scenes A and B were used for interpretation training. The range of sizes and structural differences between the POIs mean that they cannot be directly compared between scenes, even for POIs with similar classifications (e.g. buildings D and J). Features of similar size and geometries between several scenes could be compared, however, participants would be alerted to the features that they need to identify each in AOI.

Method Reflection 4: *In the Interpretation task, the POI features were a range of geometries. Some answers displayed positive or negative bias towards a correct interpretation.*

The interpretation task required an open mind because no classifications were offered as guidance, but the method of verbally reporting ideas meant some participants were reluctant to describe their interpretations, for fear of being wrong. The participants were at liberty to describe features in their own words. Those who were comfortable with talking aloud were uninhibited with their answers, whether serious or in jest, – dog, shrine, batmobile; others may have

also seen similar objects, but been reluctant about expressing their initial thoughts. The user-generated terminology used by the participants during their descriptions introduced uncertainty during the experiment, which was normally addressed within the experiment by the interviewer. This semantic uncertainty, which was identified in the analysis stage of the study, is discussed further in section 6.5.3. However, the description of features by the participants was ambiguous because the size or scale of the feature was bigger/smaller than the answer given. For example, one volunteer described POI C (hedge) as a corridor, and when prompted to re-explain reiterated the same term. When language problems arose, features could be described correctly. Occasionally, when required, Google translate (<https://translate.google.co.uk/>, accessed 01-02-14) was used to confirm words from participants' own language. Unexpectedly, this verification technique also gave more confirmation as to the participant's mental image of the POI feature, compared to a straight-forward verbal description.

6.5.2 Restriction of learning

Some qualitative results were omitted from the discussion on the basis of assumptions that were made during the method development. The participants were not told whether their measurement or interpretation answers were correct. Qualitative information was sought on the effect of dimensionality (Jones *et al.*, 2008; Seipel, 2012; Garg *et al.*, 2002) of features on 2D and 3D interpretation, so rankings were given for natural and man-made features within each of the scenes. This is based on the assumption that participants use their knowledge from the lidar overview that they were taught at the beginning of the experiment, *i.e.* natural features are volumetric and man-made are more planar. However, the participants were not made aware that, for example, POIs I and J were building points, so when asked about man-made features in Scene D during the feedback section, their answers (Appendix L) could not be used.

Method Reflection 5: *No answers were revealed to the participants during the measurement and interpretation tasks.*

One of the features in the experiment, hedge (POI C), caused confusion among the participants, who were unfamiliar with the shape. The binomial distribution analysis, which was the test for bias, is dependent on the sample of participants' proportion and number of correct answers. A sample of lidar analyst experts may result in higher likelihood of the interpretation results being correct. In this experiment, although POI G (cliff) was biased in both methods, there was enough significant difference between 2D and 3D interpretations. For lidar experts, however, an expected high bias towards correctness in both 2D and 3D may not be coupled with enough disparity between the accuracy results. This would have resulted in insignificant results, so the comparison of accuracy. In other words, the difference in accuracy (proportion of volunteers who interpreted the POI correctly) must be of sufficient significance, however, if the results for the 2D and 3D are both biased. Some training or reference features could help improve the results of both 2D and 3D interpretation methods.

The participants in this research had a range of expertise, which was randomly distributed throughout the 2D and 3D groups. Comparison between different levels of expertise could highlight categories of participants who are more affected by immersive stereoscopy. Novice users are often compared against experts (Zehner, 2010; Banic, 2014; Savage, 2006; Kastens & Ishikawa, 2006) and, providing there is a large enough pool of experts, the evaluation of 2D vs. 3D lidar visualisation could be explored.

Method Reflection 6: *The difference in 2D and 3D interpretation results was significant for the sample population.*

6.5.3 Classification Technique

The validation of feature classifications for the Interpretation task was not based on consensus, unlike Scepan *et al.* 1999 and Bicheron *et al.* 2008 in Foody *et al.* (2015), but in this case the expert interpretation did not warrant validation from other analysts. However, further verification might further increase the

confidence in the reference classification, against which all of the participant verbal classifications were compared. Work by Foody *et al.* (2013, p. 856) states that when an individual is classifying remotely-sensed data, the classifications or 'labels' from others should be known to help data accuracy assessment. If classification terms are shared between analysts/participants, e.g. vegetation, the assessment of the classifiers' accuracy can be improved because the terms are standardised. This remote-sensing research takes a non-traditional approach in the way that there is no standardised classification. By allowing the user to keep an open mind, a more investigative and exploratory approach is used. The accuracies of manual remotely-sensed data classification could be cross-analysed more easily if common classification terminology were used. During a vision experiment, using a Randot stereo test, similar to the one used for stereoacuity screening in this study, Frisby and Clatworthy (1975) forewarned 103 participants about the type of object and/or the depth at which it would appear when viewing the stereo image. This led to faster recognition of the 3D feature, attributed to 'perceptual learning', but unexpectedly, the information did not have a bearing on the performance (speed). However, this standardisation of answers can limit a human participant's cognitive assessment of the scene. The classification terms offered by the volunteers during interpretation were categorised into binary correct/incorrect results, but these could be quantified into 'fuzzy' classifications or different levels of accuracy. A more traditional remote-sensing error analysis approach could be adopted, such as a user versus producer's error matrix (Congalton, 1991; Foody, 2010), measuring the proportion of points that have been misclassified by the participants and expert interpreter respectively. Errors of commission and omission could provide an indication of those points that have been incorrectly assigned to a class (e.g. a ground point allocated to vegetation class) and those that have been omitted from a classification (e.g. a vegetation point that was allocated to something other than vegetation).

Analysis of the semantics of participant answers could add to more general remote-sensing interpretation discussions (Foody *et al.*, 2014; Comber *et al.*, 2013; Meek *et al.*, 2014). The choice of vocabulary used by participants could be inspected in the same way as Comber *et al.* (2013), who describe the

frequency of terms used in journal article abstracts over time. As opposed to recognising semantic links between publication years or topics, as Comber *et al.* (2013) do, common links might instead be established between groups of people or in response to tasks that were carried out. This could highlight influences of participant background on results and would offer more understanding into the volunteers' emotional reactions to each task. Keeping in mind that user preference should be considered alongside performance (Andre & Wickens, 1995), this might offer valuable insight into end-user adoption of remote-sensing visualisation technologies.

Method Reflection 7: *The verbal classification method generated rich qualitative data.*

6.6 Summary

For the suburban and valleyside scenes explored by participants within the experiment, the accuracy of stereo 3D feature interpretation is better than or equal to 2D. Of the ten features investigated by the participants, eight had similar 2D and 3D interpretation accuracies, but two (a chimney from Scene C and cliff-face from Scene D) had better 3D interpretations. The finding leads to the assumption that, under participant trial conditions, the extra depth perception of the 3D method gave added-value when trying to identify certain features within a lidar point cloud.

The outcome of the interpretation task unsurprisingly raises further questions that should be explored, if possible, with participants use point cloud visualisations, whether derived from lidar or other means, *e.g.* SfM. Although the results show a slight improvement in interpretation when using the 3D method, it is not clear whether the 2D method is sufficient for different applications. Owing to the location of the VR theatre used in this experiment and sample population, the trials were tailored towards a general point cloud classification task. Future similar experiments could help address this gap by tailoring the experiment towards specific users of specific point cloud datasets,

e.g. commercial production operators for urban 3D modelling or foresters using SfM data for inventory use. Further work could also borrow analysis concepts from current citizen science literature. Furthermore, GIS fuzzy literature could also be used; no levels of accuracy or fuzzy classifications were considered in the study. Interpretation answers were either deemed correct or incorrect. In addition to the quantitative results, the participant feedback revealed the observations below, and wider implications of the findings are discussed in the section 8.

- Learning process: participants apply knowledge gained from learned fundamentals of lidar and previous experience with the data.
- Verbal classification technique could be developed by using fuzzy, instead of binary correct/incorrect, accuracies and assessment of the semantics of the answers.
- Could train users to classify point clouds using interaction devices.

7.

Method Review

7. Method Review

7.1 Introduction

This chapter reviews the general methodological approach that was used for the overall experiment, in addition to those reflections within the Measurement and Interpretation chapters. Future work is also considered, in light of the methods and findings of this study. It is acknowledged that the evaluation of visualisation methods is challenging because goals are difficult to define, and it is difficult to design tests that meaningfully evaluate them (van Dam *et al.*, 2002).

7.2 Reflection on general method development (RQ3)

7.2.1 Depth perception

The study did take into account the interpupillary distance (eye separation, see section 3.2.2) of each participant, but an examination of the angle of a change in their focus may be considered intrusive. It was sufficient to ask participants if they had any eye conditions, although those with myopia, stigmatism had corrected vision. In a stereo 3D environment, it is advisable to reduce clutter in front of the user and this recommendation was followed in this study. However, for geovisualisation, the clutter (axes, North arrows, and user coordinates) provides useful information for user navigation; Zehner, *et al.* (2010) argue there is a need for 2D display (for text, maps, graphs) within or alongside virtual stereoscopic 3D environments. Kim *et al.* (1987) found that monoscopic displays are optimised with additional depth cues, *e.g.* lines of reference, which may improve performance to that similar to a stereoscopic display. This implies that the inclusion of within-scene graphical aids, such as markers or rulers, may lead to an increase in the accuracy of parameters derived from 2D lidar

visualisation, although Kim *et al* (1987) state that the stereoscopic technique had consistently lower errors. Clutter could be included in further lidar visualisation studies to explore the effects of monoscopic depth cues.

Method Reflection 8: *The study is sensitive to psychological human factors, such as stereoacuity and interpupillary distance.*

7.2.2 Human factors

Participants reportedly felt 'dozy' or were yawning by the end of the study. It was at this point, when fatigued, they were asked to recall their feelings towards and experiences of the scenes and tasks. The verbal questions were asked at the end of the session, however, similar studies could pose questions during or shortly after each task. The level of fatigue could have affected short term memory. Typically, participants were offered a break after Task 1 and although they were not forced to take this. A study into timed performance of human operators queried the 300 participants about how distracted they felt, tiredness, time of the day, and amount of coffee consumed (Van Coillie *et al.*, 2014). Although the latter may seem trivial, unmeasured external effects may have had an effect on individual performance.

Method Reflection 9: *The method mitigated against human factors, which were not measured as part of the research.*

7.2.3 Effect of pilot study

Although other visualisation studies have used completion time (Garg *et al.*, 2002; Koua *et al.*, 2006; Seipel, 2012; Van Coillie *et al.*, 2014), as a direct result of the pilot study (during which volunteers felt panicked under timed conditions) the participants were not given time constraints.

Method Reflection 10: *The pilot study was crucial for assessing task viability, although potential adaptations in experiment design should be weighed-up against the requirements of the research.*

7.2.4 Qualitative data

When designing the qualitative data collection method, it was considered that a video may be too intrusive for the interviewee. However, during the trials, volunteers communicated strongly with body language and some reacted with a surprised jolt when using the 3D glasses for the first time. While seated, one participant tried to lean around a point, illustrating the strength of the sense of presence that the 3D experience gave some users. In addition to normal gesticulations, participants got up out of their seats to explain their point on the screen. When using hand gestures in front of the screen, they experienced vergence-accommodation conflict; the person's hand appears double when focusing (and converging eyes) on the projected data (well-illustrated by Bruder *et al.*, 2013, p.10). Of the participants who carried out the full experiment, 9% further expressed themselves through drawing, either using pen and paper, or by using a mouse to highlight areas of interest on a screenshot of the point cloud.

Banic (2014) noted that, prior to interacting with 3D objects, users strategised and had an "observational/exploration step" prior to selection of points (Banic, 2014, p.16). The participants chose to ignore points that caused occlusion, and changed scale of the view. A planning step was also observed in this study, prior to carrying out the measurement task, Appendix M comments, when participants would evaluate the form of the cloud prior to/during navigation to their selected points. In the interpretation task, the users were given time to explore the scene, after pilot participants found that it was helpful to navigate and describe the scenes, for contextualisation before the POI interpretation. This familiarisation aspect was carried out in both 2D and 3D, and could be investigated in future research.

Method Reflection 11: *Rich participant feedback was generated, which could be further analysed.*

7.3 Method recommendations

The list below reiterates the recommendations that were generated from methodological development reflection (RQ3). These are collated from the measurement task, interpretation task, and the wider experiment. The recommendations influence future research, which is suggested in the next section.

- 1) Investigate the effect of different interaction devices and methods on 2D vs. 3D visualisation results.
- 2) Develop and explore comparative 2D- and 3D-appropriate measurement methods for laser-scanned vegetation.
- 3) Explore additional performance methods for 2D vs. 3D visualisation comparisons.
- 4) Randomly select a range of interpretation features for the participants to consider.
- 5) Take into account the knock-on effect of participants' performance on their ability to suitably answer the survey questions.
- 6) Users of different ability may not exhibit a significant difference in 2D vs. 3D. Design interpretation tasks around the target population.
- 7) Further explore qualitative classification methods by generating fuzzy classes and by studying the semantics of the descriptions. Alternatively, set pre-determined classifications for the participants.
- 8) Acknowledge human cognitive psychological differences as far as possible.
- 9) Either mitigate against or measure the impact of human methods.
- 10) Use a pilot study to help refine or develop tasks.
- 11) Investigate participant feedback methods if narrative to the quantitative results is desired.

7.4 Future work

This section discusses other work that could be carried out to advance aspects of this area of research.

7.4.1 2D vs. 3D evaluation

Performance indicators

When the results of 2D were the same as 3D, as was in the measurement task and some POIs in the interpretation task, more performance indicators could be used to determine any differences.

The amount of navigation carried out by participants would help to indicate whether and where motion parallax (Figure 1-3) was used by 2D and 3D groups. By moving the viewpoint side to side, points further from the view appear to move more slowly than those that are closer and give the viewer an impression of depth. This technique is used in the “wiggle” view that is available in the lidar software FUSION/LDV (McGaughey, 2014) and the constrained view that was available to participants in the 2.5D skeletal bone study carried out by Garg *et al.* (2002).

Kreylos *et al.* (2006) state that, during 2D point cloud visualisation, the requirement to move the viewpoint in order to gain the motion parallax depth cue hinders the accuracy of users’ point selection. The movement of the participants during the tasks in this study could help indicate the extent to which the volunteers still relied on this cue during the 2D and 3D methods.

According to the observations made by Kreylos *et al.* (2006), and the assumption that 3D depth perception allows greater depth (without motion parallax), participant point selection would require less movement because the stereopsis depth from binocular display would provide sufficient information. In other words, 3D could mean less navigation is required. As an alternative to assessing the 2D and 3D exploration routes, an animation could be used, which requires no interaction from the user, but creates a controlled route for all

participants. In the author's previous HMD point cloud visualisation study (Burwell *et al.*, 2012), lidar analysts state that navigation to a helpful viewpoint within the data could potentially be time-consuming during quality assurance. However, if the stereoscopic element of immersive navigation increases the likelihood of an accurate result, as Kreylos *et al.* (2006) suggest, the (re)editing could be reduced overall.

Quantitative tracking information can be enhanced by qualitative information *i.e.* colouring the tracks according to the person's feelings of risk and safety (Kwan and Ding, 2008), importance, or religious significance (Mennis *et al.*, 2013). By combining participants' virtual tracking information with associated audio commentary, a more holistic 2D vs. 3D user experience could be mapped, potentially separating performance descriptors. Participant tracks could be coloured according to their talk-aloud POI suggestions, to show development of ideas according to location, or one could map reported levels of confidence during tasks, which may vary over time. For example, during navigation, a 2D participant might understand the spatial data arrangement with increasing movement, but a 3D participant may already feel confident in their understanding of the point cloud structure because of the binocular depth cue.

Ultimately, the adoption of 3D visualisation depends on the preference of the human who is viewing the data and the employment of participants offers scope for more rigorous preference vs. performance assessment (after Andre & Wickens, 1995). The usability of interaction techniques could be assessed using the ISO 9241-9 method to compare non-keyboard input devices via 2D vs. 3D displays (Teather & Stuerzlinger, 2014).

7.4.2 2D vs. 3D human computer interaction methods

Differences in the visualisation methods are kept to a minimum during a controlled comparison (*i.e.* same interaction device, layout of GUI, *etc.*). However, this compromises the configuration of each method is not presented to the user. More method-appropriate HCI tools for 2D and 3D lidar visualisation may be required. In standard software, there is some user manipulation of data views, such as determining the size of the points, colour of

points and background, viewpoints. However, although the GUI can be maximised for each individual's personal preference, collaborative work Darken *et al.* (1999) found that participants required different orientations of virtual maps when performing search tasks and these, in turn, were dependent of the user's mental rotation abilities. This suggests that if viewpoints are taken into consideration, there is a knock-on effect to cognitive judgement.

Only two tasks were analysed in this study, but other lidar-specific applications may, or may not, benefit from 3D visualisation. The extensive review of 2D vs. stereoscopic 3D studies by McIntire *et al.* (2014) noted that 3D method was most useful for object manipulation. Alternative lidar tasks should include virtual manipulation of the dataset, including virtual tools (Gardner *et al.*, 2003; Kreylos *et al.*, 2006) that may increase on-screen monoscopic depth cues, but which would take advantage of the characteristics of 3D visualisation that are accredited in the 2D vs. 3D literature.

7.4.3 Effect of lidar data representation on 2D vs. 3D results

The lidar points used in the interpretation task were not the raw point cloud – some points were omitted, which may provide visual clues that allow human pattern recognition to fully explore a dataset. Kreylos *et al.* (2008, p.847) argue the case for 3D visual analysis of an unprocessed, raw point cloud; its data integrity expands quality control capabilities, increases the accuracy and ease of use of point selection for feature extraction, and enhances the overall analytical potential of the 3D data. (Kreylos *et al.*, 2008).

During the participant experiment, the point cloud data were the only information presented to the volunteers. Reference images were not shown to participants during the interpretation task. In a practical, real-world application of lidar analysis tasks, reference information, such as aerial photography, is often available to the end user. The addition of more contextual information may be sufficient for visual lidar analysis tasks performed adequately in 2D. However, depth perception offered by a 3D display could offer more immediate structural understanding by the user at a cognitive level that may not be met by reference to auxiliary data. Furthermore, additional context may not always

available. In the same way that Brodu and Lague (2012) visualised a worst-case representation of data (one TLS angle) for their multi-scale classification of geomorphology scenes, this study takes advantage of a worst-case scenario where the point cloud is the only information available. Furthermore, from a cognitive psychology perspective, the inclusion of several confounding variables (e.g. 2D/3D method, point colouring, reference dataset availability) leads to a result based on a group of interacting variables. The isolation of one factor that may be affecting the performance or preference of the end-user, *i.e.* visualisation method, means that the effect can be assessed independently of other variables. In stripping back the characteristics of the data to its physical, spatial position (without colour or classification), the study focuses on the physicality of the data, which is inherently three-dimensional (Warner *et al.*, 2003). However, further contextualisation of the data was considered and included in the experiment. As stated in the pilot study, section 3.3, the interpretation task data were reshown to the participants with colouring according their classification. Results were not analysed within the timeframe of the study, but once analysed would highlight whether the inclusion of automated classification information has a positive or negative effect on the participants' original (one-colour) interpretation. Additional information such as near-infrared reflectance could help classify feature types prior to visualisation. This, and other automated classification, could potentially alleviate the cognitive load of the human interpreter or, if mismatched with the point cloud, could add. In this study, as mentioned in section 3.2.1, the application of spectral colour from orthophotographs was rejected because of an offsets and lean. In future work, the addition of true-colour hues could be integrated into point clouds, especially from SfM, which is derived from true-colour photos or videos. This study presents the effect of 2D and 3D visualisation methods on lidar tasks, with a controlled representation of the data. An assessment that incorporates the inclusion or comparison of lidar with supplementary datasets may be more reflective of typical lidar task, however, user studies require control of variables. The addition of contextual information should be analysed in future studies, on the condition that variables are also individually assessed. This would require a large number of participants, to accommodate the number of tests required. It

has been predicted that citizen-science-based visual interpretation tasks are likely to become more complex (Bianchetti & MacEachren, 2015). Potentially, a large-scale citizen science study, using cheaper technologies (such as cardboard mobile-phone-enabled VR headsets, see <https://www.google.com/get/cardboard/get-cardboard/>, accessed 09-01-16) could accommodate multiple visualisation variables.

Further analysis into user navigation tracking could help explain aspects of the results 2D and 3D results. The user's position and orientation were recorded during the task, but their analysis was out of the scope of the study. Analysis of user navigation could also be used to understand how the effect of motion parallax (Figure 1-3) is exploited in 2D and 3D visualisation environments, as acknowledged by Kreylos *et al* (2006).

The study makes assumptions that the oculomotor cues in each individual were working as expected. Another aspect of the work related to quantification of stereoscopic effect vs. point cloud characteristic – does the 3D effect make it easier to see and understand structures, or the clustering of points?

7.5 Summary

By reflecting on the novel methods used during the study, examples of good practice have been identified, such as the measurement of participants' stereoacuity and the verbal discussion of point cloud classifications. The review also put forward alternative approaches that could be explored, *e.g.* the use of different interaction devices and additional performance indicators. These reflections and recommendations should be used as guidance for similar future work.

8.

Conclusion

8. Conclusion

8.1 Introduction

This research sought to evaluate whether immersive stereoscopic 3D display adds value to the analysis of lidar point cloud data, compared to a 2D visualisation. The study questions whether there are differences in the precision and/or accuracies of lidar point cloud (1) measurement and (2) interpretation. The major novelty of this research lies in its methodology, therefore, its suitability is investigated (3).

8.2 Findings

The key findings are summarised in this chapter, alongside a reminder of each hypothesis, the details of which can be referred to in Chapter 2. Further discussion on the findings can be found in the respective chapters.

8.2.1 Measurement task (RQ1)

H1 – There is a significant difference in linear measurement of lidar point cloud features derived from 2D and 3D visualisations for (a) a planar building feature (b) a volumetric vegetation feature.

The length of a roof edge was measured by volunteers and the longest diameter of a canopy of trees was also measured. Overall, it was hypothesised that there would be a significant difference between 2D and 3D measurements for planar and volumetric features. However, when comparing the frequency distributions of 2D and 3D measurement results, there was no significant difference, for both features. For the planar roof edge (a), the linear edges of the planar feature reportedly facilitated the task for both 2D and 3D participants. For both 2D and 3D methods, participants wanted to position themselves

perpendicular to the straight edge. The stereoscopic depth reportedly helped some participants understand spatial arrangement of clustered points at the roof corner, indicating that participants with good stereoacuity may actively exploit the extra depth.

The irregularity of the vegetation (b) and lack of straight edges made the measurement more challenging for both methods. In 2D, when positioned above the data, it was reported to be hard to distinguish the vegetation points, from the lower ground points. Overall, it was hypothesised that there would be a significant difference for planar and volumetric features (H1.2), but a comparison of the 2D and 3D measurement results, for both features, did not prove the hypothesis to be correct.

8.2.2 Interpretation task (RQ2)

H2 - Lidar point cloud feature interpretation is more accurate when carried out via 2D or 3D visualisation.

The interpretation task presented participants with point cloud representations of features that were structurally diverse. It was expected that there would be variation between the feature results (H2.1a), but that the 3D accuracies would be higher than 2D (H2.1b). This research shows that, in certain scenarios, interpretation of lidar point cloud features is significantly better in stereoscopic 3D, rather than 2D. Two of the ten features (chimney and cliff-face) were better interpreted by participants using the stereo method, which could be attributed to the added binocular stereo depth provided by 3D (Eysenck, 2001; Howard & Rogers, 2012; McIntire *et al.*, 2014) and insufficient pictorial depth cues in 2D. This is supportive of immersive stereo lidar literature that implies that 3D visualisation allows users to accurately interpret lidar point clouds as features (Kreylos *et al.*, 2006). Furthermore, participants remarked, as expected (H2.2), that the point cloud that represented a flat, suburban environment was easier to interpret, irrespective of method, because it was a familiar, day-to-day scene. In contrast, contextualisation of the sloped environment was challenging for both groups because of the inclined ground surface, which participants from the 2D

group (who lacked stereoscopic depth) reportedly found particularly difficult. This helps to confirm the hypothesis (H2.2) that human visualisation in both 2D and 3D is affected by the complexity of the viewed acquisition environment, similarly to automated ground-filtering lidar algorithms (Meng *et al.*, 2010). However, a higher proportion of 2D participants correctly interpreted the planar road, compared to 3D, although not at a significant level.

8.2.3 Reflection on methodological development (RQ3)

H3 - The method used for the set-up and execution of the measurement task (H3.1), interpretation (H3.2) task, and the general experiment (H3.3) is suitable, but improvements are identified.

For a 2D vs. 3D comparative experiment, which had not been previously attempted in the context of lidar, the 2D and 3D visualisations were developed to be as similar as possible, differing only in stereo depth. The experiment (H3.3) acknowledged psychological factors (notably stereoacuity and IPD) that are fundamental to the research into stereoscopic lidar visualisation. These factors had not before been explicitly incorporated into former 3D-only lidar studies (Warner *et al.*, 2003; Kreylos *et al.*, 2006; Kreylos *et al.*, 2008; Burwell *et al.*, 2012). The pilot study was crucial to developing the lidar tasks, testing the visualisation workflows, and honing subtle participant-oriented details within the experimental setting. Although human factors could be measured and analysed, this study was concerned with the effect of monocular and binocular visualisation so, instead, mitigated against the influence of human factors as far as possible. The work generated 90+ hours of audio feedback, which could be further analysed to investigate the semantics of the user-generated classifications.

With the measurement task (H3.1), the interaction design (*i.e.* gamepad device and associated code) may or may not have influenced the outcome of the measurement results, but the device used was acceptable for both 2D and 3D cognition (Mark, 1992). Additional performance methods for 2D vs. 3D visualisation comparisons, such as timed conditions, would help separate out

results where accuracy does not appear significantly different. The use of a planar and volumetric feature were prerequisites to the task, based on dimensionality concepts described by Jones *et al.* (2008), although the challenging and subjective volumetric measurement sub-task requires further consideration.

The interpretation task (H3.2) generated some results that were significantly different in 2D and 3D and those that were not. The verbal answers could be taken further by studying the semantics of the verbal classification answers, which could explore research that is currently outside the scope of this work. The employment of randomly-selected POI features could remove any bias that may be associated with the subjective selection of the POIs, which was carried out in the study.

8.3 Impact

The outcome of the study has theoretical relevance, an influence on practical application of lidar visualisations, and wider-reaching implications. Although some stereo 3D lidar visualisation products already exist on the market (McGaughey, 2014; Napier, 2011), this novel assessment of geovisualisation is, at the time of writing, the first study validating the use of immersive stereo depth for lidar tasks. The higher accuracies of 3D manual classification, compared to standard 2D, could be an indication to lidar users that the increased spatial understanding of lidar in 3D can have a positive effect on productivity. The use of stereo 3D during lidar quality checks is likely to increase the accuracy of lidar-derived products, such as maps, digital terrain models (DTMs) and aerial imagery. The 3D technique used in this study did not generate significantly better results than the standard 2D approach, when applied to linear measurements of visualised data. In practical terms, users could sufficiently extract linear lengths from point clouds in 2D software such as commercial Terrasolid (Terrasolid, 2013a), open-source CloudCompare (Girardeau-Montaut, 2014), and others.

Given the recent proliferation of virtual reality headsets in the gaming industry, there is a renewed general interest in VR and 3D visualisation technologies. This, in turn, could see an increase in the use of stereoscopic viewing for lidar point clouds and similar three-dimensional clouds of information. Further development in this area demands collaboration between the remote-sensing community, technology developers, human-computer interaction experts, and cognitive psychologists. This will also lead to advancements in this interdisciplinary area, including the investigation into other 3D data tasks. Outside the field of remote-sensing, the findings are relevant to medical applications. Garg *et al.* (2002) and Seipel *et al.* (2012) suggested analysis of 2D vs. 3D visualisation of volumetric features, as opposed to planar features. This was addressed through the study, at a geographical scale, by using both volumetric and planar features. Here, acquisition, processing, and analysis are discussed in the context of lidar. However, the study did not permit reference to any auxiliary information, *e.g.* aerial photographs; if more context were available, 2D lidar analysis could be adequate for the user's needs. Future work could investigate the utility of the 2D and 3D methods when coupled with contextual information, to help determine if their adoption is warranted for certain tasks. Furthermore, other point clouds, *e.g.* SfM, and 3D data forms may benefit from the same principles of this lidar-oriented study and aspects of the research approach can be applied elsewhere.

8.4 Conclusion

Three-dimensional scans of the environment, which are used for 3D mapping, are generally visualised on-screen in 2D. As a result, the depth of the spatial information is lost, potentially along with our maximal spatial understanding of the data. This research sought to compare 2D visualisation with stereoscopic 3D to help evaluate whether a true-3D representation of three-dimensional lidar datasets adds more value than 2D. Prior to this research, no other comparative study had been performed to investigate this gap in the literature. Novel experiments were developed to simulate manual lidar analysis tasks in a

controlled setting and a visualisation system was developed to allow participants to view point clouds in 2D or stereo 3D. The outcome of the study suggests that stereo 3D visualisation does add to our visual analysis of point cloud features for certain tasks and the following general conclusions were drawn from the experiment analysis:

- The generation of linear measurement vectors along planar surfaces and through volumetric forms is statistically the same in 2D and stereo 3D visualisation.
- The stereoscopic 3D interpretation of features from point clouds is worthwhile because its accuracy is either the same as or better than 2D.
- The novel methods in the experiment were suitably developed and further recommendations are suggested.

The findings suggest that, for ALS systems, there would be an increase in accuracies of lidar-derived products if immersive 3D visualisation were applied to manual lidar point cloud classification. However, 2D visualisation may be sufficient for the manual estimation of lengths from point clouds, although investigation into other performance factors may contest the outcome of the performed measurement task. Overall, the research highlights that the true visual-analytical capability of human interpreters is not fully exploited during the 2D interpretation of lidar point clouds. Stereoscopic vision can be unlocked by using 3D visualisation technologies, allowing spatial understanding and deeper analysis of discrete lidar data. The development and implementation of this research has demonstrated that 3D visualisation of lidar point clouds does add value to 2D for certain tasks.

B.

Bibliography

Bibliography

- Airbus Defence and Space Ltd., 2013a. Discrete airborne lidar dataset for 1kmsq area of Bristol, UK (Ordnance Survey Grid Refs: ST565725 and ST575725) in binary format (.bin) acquired 25-Nov-2008. Leicestershire, UK: © Airbus Defence and Space Ltd.
- Airbus Defence and Space Ltd., 2013b. Discrete airborne lidar dataset for 200msq area of London, UK (Ordnance Survey Grid Ref: TQ240800) in binary format (.bin) acquired 04-Jul-2008. Leicestershire, UK: © Airbus Defence and Space Ltd.
- Anderson, K. and Gaston, K.J., 2013. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*. 11, 138-146.
- Andre, A.D. and Wickens, C.D., 1995. When users want what's not best for them. *Ergonomics in Design: The Quarterly of Human Factors Applications*. 3, 10-14.
- Baltsavias, E.P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry and Remote Sensing*. 58, 129-151.
- Banic, A., 2014. Selection classification for interaction with immersive volumetric visualizations. In Yamamoto, S., ed, *Human interface and the management of information. information and knowledge design and evaluation: 16th international conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings, Part I.*, LNCS8521, Switzerland: Springer International Publishing. 10-21.
- BBC, 2013. BBC 3D programming 'on hold' indefinitely. *BBC News*. 5 July 2013 <http://www.bbc.co.uk/news/entertainment-arts-23195479> (accessed 08/07/13)
- Bentley Systems Inc., 2010. MicroStation V8i - 3D CAD design and modeling software for architecture, engineering, construction and operations. v8i (SELECTseries 2), PA, USA: Bentley Systems Inc.
- Bernardin, T., Cowgill, E., Kreylos, O., Bowles, C., Gold, P., Hamann, B., and Kellogg, L., 2011. Crusta: A new virtual globe for real-time visualization of sub-meter digital topography at planetary scales. *Computers and Geosciences*. 37, 75-85.

- Bianchetti, R.A. and MacEachren, A.M., 2015. Cognitive themes emerging from air photo interpretation texts published to 1960. *ISPRS International Journal of Geo-Information*. 4, 551-571.
- Bleisch, S. and Dykes, J., 2014. Quantitative data graphics in 3D desktop-based virtual environments – an evaluation. *International Journal of Digital Earth*. 1-17.
- Brodlie, K., Dykes, D., Gillings, M., Haklay, M.E., Kitchin, R., and Kraak, M.J., 2002. Geography in VR: Context. In Fisher, P. and Unwin, D., eds, *Virtual Reality in Geography*. London: Taylor & Francis. 7-16.
- Brodu, N. and Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing*. 68, 121-134.
- Bruder, G., Steinicke, F. and Sturzlinger, W., 2013. To touch or not to touch?: comparing 2d touch and 3d mid-air interaction on stereoscopic tabletop surfaces, *1st Symposium on Spatial User Interaction*, 20-21 July 2013, ACM 9-16.
- Buckley, S.J., Howell, J.A., Enge, H.D., and Kurz, T.H., 2008. Terrestrial laser scanning in geology: Data acquisition, processing and accuracy considerations. *Journal of the Geological Society*. 165, 625-638.
- Burwell, C., Jarvis, C., and Tansey, K., 2012. The potential for using 3D visualization for data exploration, error correction and analysis of LiDAR point clouds. *Remote Sensing Letters*. 3, 481-490.
- Carr, K. and England, R., eds, 1995. *Simulated and virtual realities: Elements of perception*. London: Taylor & Francis.
- Chen, L., Wei, H., and Ferryman, J., 2013. A survey of human motion analysis using depth imagery. *Pattern Recognition Letters*. 34, 1995-2006.
- Comber, A., See, L., Fritz, S., Van der Velde, M., Perger, C., and Foody, G., 2013. Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied Earth Observation and Geoinformation*. 23, 37-48.
- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*. 37, 35-46.
- Danson, F.M., Hetherington, D., Morsdorf, F., Koetz, B., and Allgower, B., 2007. Forest canopy gap fraction from terrestrial laser scanning. *IEEE Geoscience and Remote Sensing Letters*. 4, 157-160.
- Darken, R.P. and Cevik, H., 1999. Map usage in virtual environments: Orientation issues, *IEEE Virtual Reality*, 13 - 19 March 1999, IEEE, 133-140.

- Döllner, J. and Hinrichs, K., 2000. An object-oriented approach for integrating 3D visualization systems and GIS. *Computers & Geosciences*. 26, 67-76.
- Dytham, C., 2006. *Choosing and using statistics: A biologist's guide*. 2nd ed., Oxford, UK: Blackwell Publishing.
- Edge®, 2013. The real thing: Oculus Rift is a bold bid to bring back VR. if it succeeds, gaming will never be the same again. *Edge Magazine*, 254, 74-80.
- Edler, D., Huber, O., Knust, C., Buchroithner, M.F., and Dickmann, F., 2014. Spreading map information over different depth layers—an improvement for map-reading efficiency? *Cartographica: The International Journal for Geographic Information and Geovisualization*. 49, 153-163.
- ESRI®, 2014. ArcGIS 10.2.2 for desktop. 10.2.2.3552 ed., Redlands, CA: Environmental Systems Research Institute (ESRI).
- ESRI®, 2013. *ArcGIS Help 10.1: An overview of editing LAS datasets*. <http://resources.arcgis.com/en/help/main/10.1/015w/015w0000003z000000.htm> (accessed 08/04/2015).
- Eysenck, M.W., 2001. *Cognitive psychology: A student's handbook*. 4th ed., Somerset, UK: Taylor & Francis.
- Faust, N.L., 1995. The Virtual-Reality of GIS. *Environment and Planning B-Planning & Design*. 22, 257-268.
- Feng, D., Marshburn, D., Jen, D., Weinberg, R.J., Taylor II, R.M., and Burette, A., 2007. Stepping into the third dimension. *The Journal of Neuroscience*. 27, 12757-12760.
- Flood, M., 2001. Lidar activities and research priorities in the commercial sector. *International Archives of Photogrammetry and Remote Sensing*. 34, 3-7.
- Foody, G.M., See, L., Fritz, S., van der Velde, M., Perger, C., Schill, C., Boyd, D.S., and Comber, A., 2015. Accurate attribute mapping from volunteered geographic information: issues of volunteer quantity and quality. *The Cartographic Journal*. 52(4), 336-344.
- Foody, G.M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., and Boyd, D.S., 2013. Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project. *Transactions in GIS*. 17, 847-860.
- Foody, G.M., 2010. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*. 114, 2271-2285.

- Frisby, J.P. and Clatworthy, J.L., 1975. Learning to see complex random-dot stereograms. *Perception*. 4, 173-178.
- Fuhrmann, S., Ahonen-Rainio, P., Edsall, R.M., Fabrikant, S.I., Koua, E.L., Tobon, C., Ware, C. and Wilson, S., 2005. Making useful and useable geovisualization: Design and evaluation issues. In Dykes, J., MacEachren, A. M. and Kraak, M. J., eds, *Exploring Geovisualization*. First ed., Amsterdam: Elsevier. 553-566.
- Gahegan, M., 1999. Four barriers to the development of effective exploratory visualisation tools for the geosciences. *International Journal of Geographical Information Science*. 13, 289-309.
- Gardin, S., van Laere, S.M., van Coillie, F.M., Anseel, F., Duyck, W., de Wulf, R.R., and Verbeke, L.P., 2011. Remote sensing meets psychology: a concept for operator performance assessment. *Remote Sensing Letters*. 2, 251-257.
- Gardner, J.V., Warner, T., Nellis, M.D., Brandtberg, T., 2003. Virtual reality technology for lidar data analysis, In N.L. Faust and W.E. Roper, eds., *Geo-Spatial and Temporal Image and Data Exploitation III*, 21 - 24 April 2003, The International Society for Optical Engineering (SPIE) 48-57.
- Garg, A.X., Norman, G.R., Eva, K.W., Spero, L., and Sharan, S., 2002. Is there any real virtue of virtual reality?: The minor role of multiple orientations in learning anatomy from computers. *Academic Medicine*. 77, S97-S99.
- Garrett, B.L., 2014. Undertaking recreational trespass: urban exploration and infiltration. *Transactions of the Institute of British Geographers*. 39, 1-13.
- GeoPerspectives, 2013a. RGB aerial photography for Bristol region ST5672, captured 14-04-2007. GEOSTORE_ST5672_2007-04-14, Leicestershire, UK: © GeoPerspectives.
- GeoPerspectives, 2013b. RGB aerial photography for London region TQ2480, captured 24-06-2010. GEOSTORE_TQ2480_2010-06-27, Leicestershire, UK: © GeoPerspectives.
- Gertman, V., Olsoy, P., Glenn, N., and Joshi, A., 2012. RSVP: Remote Sensing Visualization Platform for data fusion, *IEEE Virtual Reality Workshop on Immersive Visualization 2012*, 4-8 March 2012, IEEE.
- Gibson, J.J., 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gigli, G. and Casagli, N., 2011. Semi-automatic extraction of rock mass structural data from high resolution LIDAR point clouds. *International Journal of Rock Mechanics and Mining Sciences*. 48, 187-198.

- Gilbert, R.O., 1987. *Statistical methods for environmental pollution monitoring*. New York: Van Nostrand Reinhold.
- Girardeau-Montaut, D., 2014. CloudCompare: 3D point cloud and mesh processing software open source project. version: 2.5.5.2 [Windows 64-bit], Release date: 21-07-14, Clamart, France <http://www.cloudcompare.org/>: EDF R&D, Telecom ParisTech.
- Gong, P., Li, Z., Huang, H., Sun, G., and Wang, L., 2011. ICESat GLAS data for urban environment monitoring. *IEEE Transactions on Geoscience and Remote Sensing*. 49, 1158-1172.
- Google, 2014. Street View - Google Maps. CA, USA: © Google.
- Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environments. *Journal of Photogrammetry and Remote Sensing*. 54, 130-137.
- Haala, N. and Kada, M., 2010. An update on automatic 3D building reconstruction. *Journal of Photogrammetry and Remote Sensing*. 65, 570-580.
- Haklay, M.E., 2002. Virtual Reality and GIS: Applications, trends and directions. In Fisher, P. and Unwin, D., eds, *Virtual Reality in Geography*. London: Taylor & Francis. 47-57.
- Harris, J.M., 2014. Volume perception: Disparity extraction and depth representation in complex three-dimensional environments. *Journal of Vision*. 14, 11.
- Hegarty, M., Smallman, H.S., Stull, A.T., and Canham, M.S., 2009. Naive cartography: How intuitions about display configuration can hurt performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*. 44, 171-186.
- Helmut, M., 2008. Object extraction in photogrammetric computer vision. *ISPRS Journal of Photogrammetry and Remote Sensing*. 63, 213-222.
- Hertzum, M. and Jacobsen, N.E., 2001. The evaluator effect: a chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*. 13, 421-443.
- Hillaire, S., Lecuyer, A., Cozot, R., and Casiez, G., 2008. Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments, In Lin, M., Stted, A., and Cruz-Neira, C., eds. *Virtual Reality Conference*, 8-12 March 2008, IEEE. 47-50.
- Himmelsbach, M., Hundelshausen, F.V. and Wuensche, H., 2010. Fast segmentation of 3D point clouds for ground vehicles, *Intelligent Vehicles Symposium (IV)*, 21-24 June 2010, IEEE 560-565.

- Hoang, R.V., Sgambati, M.R., Brown, T.J., Coming, D.S., and Harris Jr., F.C., 2010. VFire: Immersive wildfire simulation and visualization. *Computers & Graphics*. 34, 655-664.
- Hodza, P., 2009. Evaluating user experience of experiential GIS. *Transactions in GIS*. 13, 503-525.
- Hoffman, D.M., Girshick, A.R., Akeley, K., and Banks, M.S., 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*. 8, 1-30.
- Holmgren, J. and Persson, Á, 2004. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*. 90, 415–423.
- Howard, I.P. and Rogers, B.J., 2012. *Perceiving in Depth, Volume 2: Stereoscopic Vision*. Shanghai, China: Oxford University Press, Inc.
- Isenburg, M., 2013. LASTools - efficient tools for LiDAR processing (with LASzip). Version 06-05-13, Gilching, Germany: rapidlasso GmbH.
- Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., and Fitzgibbon, A., 2011. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera, *24th Annual ACM Symposium on User Interface Software and Technology*, 16 - 19 October 2011, ACM 559-568.
- Jáuregui, D.A.G., Argelaguet, F., and Lecuyer, A., 2012. Design and evaluation of 3D cursors and motion parallax for the exploration of desktop virtual environments, In Billinghamurst, M., LaViola, J., and Lecuyer, A., eds., *IEEE Symposium on 3D User Interfaces (3DUI) 2012*, 4-5 March 2012 2012, IEEE. 69-76.
- Jones, R.R., Wawrzyniec, T.F., Holliman, N.S., McCaffrey, K.J.W., Imber, J., and Holdsworth, R.E., 2008. Describing the dimensionality of geospatial data in the earth sciences—recommendations for nomenclature. *Geosphere*. 4, 354-359.
- Jones, R.R., McCaffrey, K.J.W., Clegg, P., Wilson, R.W., Holliman, N.S., Holdsworth, R.E., Imber, J., and Waggott, S., 2009. Integration of regional to outcrop digital data: 3D visualization of multi-scale geological models. *Computers & Geosciences*. 35, 4-18.
- Kastens, K.A. and Ishikawa, T., 2006. Spatial thinking in the geosciences and cognitive sciences. In Manduca, C. A. and Mogk, D. W., eds, *Earth and Mind: How Geologists Think and Learn About the Earth*. Colorado, USA: Geological Society of America, Incorporated. 53-76.
- Kato, A., Obanawa, H., Hayakawa, Y., Watanabe, M., Yamaguchi, Y., and Enoki, T., 2015. Fusion between UAV-SFM and terrestrial laser scanner for field

- validation of satellite remote sensing, *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, 2642-2645.
- Kim, W.S., Ellis, S.R., Tyler, M.E., Hannaford, B., and Stark, L.W., 1987. Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. *IEEE Transactions On Systems, Man and Cybernetics*. 17, 61-72.
- Koua, E., Maceachren, A., and Kraak, M., 2006. Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographical Information Science*. 20, 425-448.
- Kovač, B. and Žalik, B., 2010. Visualization of LIDAR datasets using point-based rendering technique. *Computers & Geosciences*. 36, 1443-1450.
- Kraak, M.J., 1998. The cartographic visualization process: from presentation to exploration. *The Cartographic Journal*. 35, 11-15.
- Kraak, M.J., 1993. Three-dimensional map design. *The Cartographic Journal*. 30, 188-194.
- Kreylos, O., Bawden, G., Bernardin, T., Billen, M.I., Cowgill, E.S., Gold, R.D., Hamann, B., Jadamec, M., Kellogg, L.H., Staadt, O.G., and Sumner, D.Y., 2006. Enabling scientific workflows in virtual reality, *International Conference on Virtual Reality Continuum and its Applications (VRCIA)*, 14 - 17 June 2006, 155-162.
- Kreylos, O., W. Bawden, G., and Kellogg, L.H., 2008. Immersive visualization and analysis of LiDAR data. *Lecture Notes in Computer Science*. 5358, 846-855.
- Lai, P.C., Kwong, K., and Mak, A.S.H., 2010. Assessing the applicability and effectiveness of 3D visualisation in environmental impact assessment. *Environment and Planning B - Planning & Design*. 37, 221-233.
- Lin, H. and Batty, M., 2009. Virtual Geographic Environments: A Primer. In Lin, H. and Batty, M., eds, *Virtual Geographic Environments*. Beijing: Science Press. 1-10.
- MacEachren, A.M. and Kraak, M., 2001. Research challenges in geovisualization. *Cartography and Geographic Information Science*. 28, 3-12.
- Mancera-Taboada, J., Rodriguez-González, P., González-Aguilera, D., Finat, J., San José, J., Fernández, Juan J., Martínez, J., and Martínez, R., 2011. From the point cloud to virtual and augmented reality: digital accessibility for disabled people in San Martin's Church (Segovia) and its surroundings, *The 2011 International Conference on Computational Science and its Applications - Part II*, 20-23 June, 2011 2011, Springer-Verlag, 303-317.

- Mann, H.B. and Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*. 18, 50-60.
- Mark, D.M., 1992. Spatial metaphors for human-computer interaction, *Fifth International Symposium on Spatial Data Handling*, August 1992 1992, 104-112.
- McAuliffe, M.J., Lalonde, F.M., McGarry, D., Gandler, W., Csaky, K., and Trus, B.L. 2001. Medical image processing, analysis and visualization in clinical research, *14th IEEE Symposium on Computer-Based Medical Systems (CBMS)*, 26 - 27 July 2001, 381-388.
- McCaffrey, K.W., Jones, R., Holdsworth, R., Wilson, R., Clegg, P., Imber, J., Holliman, N., and Trinks, I., 2005. Unlocking the spatial dimension: digital technologies and the future of geoscience fieldwork.. *Journal of the Geological Society*. 162, 927-938.
- McCall, C., 2008. *Vizard teacher in a book*. July 2008 [http://www.virtual3d.co.kr/download/files/Vizard Teacher in a Book July 2008.pdf](http://www.virtual3d.co.kr/download/files/Vizard_Teacher_in_a_Book_July_2008.pdf) (accessed 08/08/14) Santa Barbara, CA.: Worldviz.
- McGaughey, R.J., 2014. FUSION/LDV: Software for lidar data analysis and visualization. version 3.42 [not using LASzip.dll], Build date 28-03-14, Seattle, WA, USA: Forest Service, Pacific Northwest Research Station, U.S. Department of Agriculture (USDA),
- McIntire, J.P., Havig, P.R., and Geiselman, E.E., 2014. Stereoscopic 3D displays and human performance: A comprehensive review. *Displays*. 35, 18-26.
- Meek, S., Jackson, M.J., Leibovici, D.G., 2014. A flexible framework for assessing the quality of crowdsourced data.
- Meng, X., Currit, N., and Zhao, K., 2010. Ground filtering algorithms for airborne lidar data: a review of critical issues. *Remote Sensing*. 2, 833-860.
- Mennis, J., Mason, M.J., and Cao, Y., 2013. Qualitative GIS and the visualization of narrative activity space data. *International Journal of Geographical Information Science*. 27, 267-291.
- Microsoft Corporation, 2015. Kinect for Microsoft.
- Microsoft Corporation, 2012a. Microsoft® Excel® 2013: Part of Microsoft Office Home and Student 2013. 15.0.4667.1000, WA, USA: Microsoft Corporation.
- Microsoft Corporation, 2012b. Microsoft® Word® 2013: Part of Microsoft Office Home and Student 2013. 15.0.4667.1000, WA, USA: Microsoft Corporation.

- Mitasova, H., Harmon, R.S., Weaver, K.J., Lyons, N.J., and Overton, M.F., 2012. Scientific visualization of landscapes and landforms. *Geomorphology*. 137, 122-137.
- Myers, J.L., Well, A. and Lorch, R.F., Jr, 2010. *Research design and statistical analysis*. 3rd ed., Hove, Great Britain: Routledge.
- Napier, B., 2011. GeoVisionary: virtual fieldwork for real geologists. Feb 2011, Sensors And Systems.
- NCH Software, 2013. Express Scribe Pro transcription software. v5.63 (Windows), Canberra, Australia: NCH Software Pty Ltd.
- Nebiker, S., Bleisch, S., and Christen, M., 2010. Rich point clouds in virtual globes - A new paradigm in city modeling?. *Computers, Environment and Urban Systems*. 34, 508-517.
- Neves, N., Silva, J.P., Gonçalves, P., Muchaxo, J., Silva, J.M., and Câmara, A., 1997. Cognitive spaces and metaphors: A solution for interacting with spatial data. *Computers & Geosciences*. 23, 483-488.
- Oculus VR, 2015a. The all new Oculus Rift development kit 2 (DK2) - Virtual Reality Headset. www.oculus.com/dk2/ (accessed 05/09/15): Oculus VR, LLC.
- Oculus VR, 2015b. Oculus Rift developer kit 1 (DK1) - Virtual Reality Headset for Immersive 3D Gaming. www.oculus.com/rift/ (accessed 05/09/15): Oculus VR, LCC.
- Omasa, K., Hosoi, F., Uenishi, T.M., Shimizu, Y., and Akiyama, Y., 2008. Three-dimensional modeling of an urban park and trees by combined airborne and portable on-ground scanning LIDAR remote sensing. *Environmental Modeling and Assessment*. 13, 473-481.
- Omasa, K., Hosoi, F., and Konishi, A., 2007. 3D lidar imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany*. 58, 881-898.
- Optech, 2008. *ALTM Gemini summary specification sheet*. Canada (www.optech.ca): Optech Incorporated.
- Park, H.S., Lee, H.M., Adeli, H., and Lee, I., 2007. A new approach for health monitoring of structures: Terrestrial laser scanning. *Computer-Aided Civil and Infrastructure Engineering*. 22, 19-30.
- Piryankova, I.V., de la Rosa, S., Kloos, U., Bühlhoff, H.H., and Mohler, B.J., 2013. Egocentric distance perception in large screen immersive displays. *Displays*. 34, 153-164.

- Preece, J., Sharp, S. and Rogers, Y., 2007. *Interaction Design: Beyond Human-Computer Interaction*. 2nd ed., Chichester: John Wiley & Sons Ltd.
- Puttonen, E., Jaakkola, A., Litkey, P., and Hyyppä, J., 2011. Tree classification with fused mobile laser scanning and hyperspectral data. *Sensors*. 11, 5158-5182.
- QSR International, 2013. NVIVO 10 for Windows. 10.0.360.0 SP3 (32-bit), Victoria, Australia: QSR International Pty Ltd.
- R Core Team, 2015. *R Documentation for package 'stats' version 2.15.3 - Exact Binomial Test*. <https://stat.ethz.ch/R-manual/R-patched/library/stats/html/binom.test.html> (accessed 04/07/15).
- R Core Team, 2014. R: A language and environment for statistical computing. R-3.1.2 for Windows (32/64 bit), Vienna, Austria: R Foundation for Statistical Computing <http://www.R-project.org> (accessed 04/07/15).
- Raber, G.T., Jensen, J.R., Schill, S.R., and Schuckman, K., 2002. Creation of digital terrain models using an adaptive lidar vegetation point removal process. *Photogrammetric Engineering and Remote Sensing*. 68, 1307-1314.
- Richter, R. and Döllner, J., 2010. Out-of-core real-time visualization of massive 3D point clouds, *7th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH '10)*, 21 - 23 June 2010, 121-128.
- Robinson, A.C., 2013. What is geographic expertise? Does it matter?, *Computer-Human Interaction (CHI) 2013 - Changing Perspectives*, April 27 - 28 2013, Computer-Human Interaction (CHI) 2013.
- Roth, R.E., 2015. Interactive maps: What we know and what we need to know. *Journal of Spatial Information Science*. 59-115.
- Royston, P., 1995. Remark AS R94: A remark on algorithm AS 181: The W-test for normality. *Applied Statistics*. 547-551.
- Saitek, 2007. P2900 Wireless PC Gamepad. Hong Kong, China: Saitek Ltd.
- Sanz-Cortiella, R., Llorens-Calveras, J., Escolà, A., Arnó-Satorra, J., Ribes-Dasi, M., Masip-Vilalta, J., Camp, F., Gràcia-Aguilá, F., Solanelles-Batlle, F., Planas-DeMartí, S., Pallejà-Cabré, T., Palacin-Roca, J., Gregorio-Lopez, E., Del-Moral-Martínez, I., and Rosell-Polo, J.R., 2011. Innovative LIDAR 3D dynamic measurement system to estimate fruit-tree leaf area. *Sensors*. 11, 5769-5791.
- Savage, D.M., 2006. *The advantages and disadvantages of three-dimensional maps for focused and integrative map analysis performance by novice and experienced users*. Ph. D. North Carolina State University (NCSU)

- Institutional Repository. <http://www.lib.ncsu.edu/resolver/1840.16/3546> (accessed 02/04/14)
- Seipel, S., 2012. Evaluating 2D and 3D geovisualisations for basic spatial assessment. *Behaviour & Information Technology*. 1-14.
- Sgambati, M., Koepnick, S., Coming, D., Lancaster, N., and Harris, F., 2011. Immersive visualization and interactive analysis of ground penetrating radar data. 6939, 33-44.
- Shapiro, S.S. and Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). *Biometrika*. 591-611.
- Sithole, G. and Vosselman, G., 2005. Filtering of airborne laser scanner data based on segmented point clouds. WG III/3 III/4, Enschede, The Netherlands: ISPRS Journal of Photogrammetry and Remote Sensing Workshop.
- Slocum, T.A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D.R., Fuhrmann, S., and Hedley, N.R., 2001. Cognitive and usability issues in geovisualization. *Cartography and Geographic Information Science*. 28, 61-75.
- Smallman, H.S., St. John, M., Oonk, H.M., and Cowen, M.B., 2001. Information availability in 2D and 3D displays. *Computer Graphics and Applications, IEEE*. 21, 51-57.
- Söderman, U., Persson, Å, Töpel, J., and Ahlberg, S., 2005. On analysis and visualization of full-waveform airborne laser scanner data. *Proceedings of SPIE - the International Society for Optical Engineering*. 5791, 184-192.
- Soininen, A., 2005. TerraScan User's Guide. Edition 1998-2005 (Published 03-10-2005) Helsinki, Finland: Terrasolid Limited.
- Stereo Optical Co., 2009. Randot™ Stereotest. Chicago, USA: Stereo Optical Co., Inc.
- Suomalainen, J., Hakala, T., Kaartinen, H., Räikkönen, E., and Kaasalainen, S., 2011. Demonstration of a virtual active hyperspectral LiDAR in automated point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*. 66, 637-641.
- Sutherland, I., 1968. A head-mounted three dimensional display, *Joint Computer Conference, Part I*, 9-11 December 1968, ACM, 757-764.
- Sweney, M., 2015. Sky to shut its 3D TV channel. *The Guardian*. 28 April 2015 <http://www.theguardian.com/media/2015/apr/28/sky-to-shut-its-3d-tv-channel> (accessed 30/04/15)

- Teather, R.J. and Stuerzlinger, W., 2014. Visual aids in 3d point selection experiments, *The 2nd ACM Symposium on Spatial User Interaction*, 4-5 October 2014, ACM, 127-136.
- Terrasolid, 2013a. TerraScan – Software for LiDAR data processing and 3d vector data creation. v013.004, Helsinki, Finland: Terrasolid.
- Terrasolid, 2013b. Terrasolid - Software for lidar processing. v13, Helsinki, Finland: Terrasolid.
- Thomsen, J.S., Laib, A., Koller, B., Prohaska, S., Mosekilde, L., and Gowin, W., 2005. Stereological measures of trabecular bone structure: comparison of 3D micro computed tomography with 2D histological sections in human proximal tibial bone biopsies. *Journal of Microscopy*. 218, 171-179.
- Tovée, M.J., 1996. *An introduction to the visual system*. First ed., Cambridge, UK: Cambridge University Press.
- Trinks, I., Clegg, P., McCaffrey, K., Jones, R., Hobbs, R., Holdsworth, B., Holliman, N., Imber, J., Waggott, S., and Wilson, R., 2005. Mapping and analysing virtual outcrops. *Visual Geosciences*. 10, 13-19.
- Van Coillie, F.M., Gardin, S., Anseel, F., Duyck, W., Verbeke, L.P., and De Wulf, R.R., 2014. Variability of operator performance in remote-sensing image interpretation: the importance of human and external factors. *International Journal of Remote Sensing*. 35, 754-778.
- van Dam, A., Laidlaw, D.H., and Simpson, R.M., 2002. Experiments in immersive virtual reality for scientific visualization. *Computers & Graphics*. 26, 535-555.
- Van Leeuwen, M., Hilker, T., Coops, N.C., Frazer, G., Wulder, M.A., Newnham, G.J., and Culvenor, D.S., 2011. Assessment of standing wood and fiber quality using ground and airborne laser scanning: A review. *Forest Ecology and Management*. 261, 1467-1478.
- Vince, J., 1998. *Essential Virtual Reality Fast: How to Understand the Techniques of Virtual Reality*. First ed., London: Springer.
- Walter, T., Shattuck, D.W., Baldock, R., Bastin, M.E., Carpenter, A.E., Duce, S., Ellenberg, J., Fraser, A., Hamilton, N., Pieper, S., Ragan, M.A., Schneider, J.E., Tomancak, P., and Hériché, J., 2010. Visualization of image data from cells to organisms. *Nature Methods*. 7, S26-S41.
- Ware, C., 2012. *Information visualization: perception for design*. 3rd ed., MA, USA: Elsevier.

- Warner, T., Nellis, M.D., Brandtberg, T., McGraw, J.B., and Gardner, J.V., 2003. The potential of virtual reality technology for analysis of remotely sensed data: A LiDAR case study. *Geocarto International*. 18, 25-32.
- Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., and Reynolds, J.M., 2012. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*. 179, 300-314.
- Wilcox, L.M. and Duke, P.A., 2005. Spatial and temporal properties of stereoscopic surface interpolation. *Perception*. 34, 1325-1338.
- WorldViz, 2010a. Vizard Lite: Virtual Reality software toolkit. Lite edition, version 3.18.0002, Santa Barbara, CA: WorldViz LCC.
- WorldViz, 2010b. Vizard: Virtual Reality software toolkit. Development edition, version 3.18.0002, Santa Barbara, CA: WorldViz LCC.
- Yan, W.Y., Shaker, A., and El-Ashmawy, N., 2015. Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment*. 158, 295-310.
- Yang, X., Strahler, A.H., Schaaf, C.B., Jupp, D.L.B., Yao, T., Zhao, F., Wang, Z., Culvenor, D.S., Newnham, G.J., Lovell, J.L., Dubayah, R.O., Woodcock, C.E., and Ni-Meister, W., 2013. Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). *Remote Sensing of Environment*. 135, 36-51.
- Yates, F., 1934. Contingency tables involving small numbers and the chi-square test. *Journal of the Royal Statistical Society*. 1, 217-235.
- Zaroff, C.M., Knutelska, M., and Frumkes, T.E., 2003. Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. *Investigative Ophthalmology & Visual Science*. 44, 891-900.
- Zehner, B., 2010. Mixing Virtual Reality and 2D visualization: using virtual environments as visual 3d information systems for discussion of data from geo- and environmental sciences, *International Conference on Computer Graphics Theory and Applications (GRAPP)*, 17-21 May 2010, 364-369.
- Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., and Gerard, P.D., 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment*. 87, 171-182.
- Zwally, H.J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., Bufton, J., Dezio, J., Hancock, D., Harding, D., Herring, T., Minster, B., Quinn, K., Palm, S., Spinhirne, J., and Thomas, R., 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics*. 34, 405-445.

A.

Appendices

List of appendices

| | |
|--|---------------|
| List of appendices | 221 |
| APPENDIX A: LIDAR DATA PROCESSING | - 1 - |
| AIRBORNE INSTRUMENT SPECIFICATION | - 1 - |
| SCREENSHOT OF LONDON LIDAR DATA ISSUE..... | - 1 - |
| LASTOOLS SCRIPT | - 2 - |
| FLIGHTLINE SCREENSHOTS..... | - 3 - |
| DENSITY SCREENSHOTS..... | - 4 - |
| APPENDIX B: VISUALISATION CODE | - 5 - |
| LIST OF MAIN MODULES | - 5 - |
| LIST OF MODULES CALLED INTO MAIN MODULES..... | - 5 - |
| MAIN MODULE (.PY) | - 6 - |
| APPENDIX C: VISUALISED DATASETS | - 13 - |
| SCENE A – POINT CLOUD DATASET..... | - 13 - |
| SCENE A – REFERENCE PHOTOS SHOWN TO PARTICIPANTS | - 18 - |
| SCENE B– POINT CLOUD DATASET | - 20 - |
| SCENE B – REFERENCE PHOTOS SHOWN TO PARTICIPANTS | - 25 - |
| SCENE C – POINT CLOUD DATASET | - 27 - |
| SCENE D – POINT CLOUD DATASET | - 32 - |
| APPENDIX D: LESSONS LEARNT FROM PILOT | - 37 - |
| APPENDIX E: DISPLAY METHOD PER SCENE | - 38 - |
| METHOD USED BY EACH PARTICIPANT FOR SCENES A-D..... | - 38 - |
| APPENDIX F: FLOW DIAGRAM OF PARTICIPANT EXPERIMENT | - 39 - |
| APPENDIX G: PARTICIPANT DOCUMENTS | - 40 - |
| INFORMATION SHEET..... | - 40 - |
| CONSENT FORM | - 42 - |
| APPENDIX H: INTERVIEWER SCRIPT | - 43 - |
| 1. INTRODUCTION | - 43 - |
| 2. MEASUREMENT TASK..... | - 50 - |
| 3. INTERPRETATION TASK..... | - 53 - |
| 4. FEEDBACK QUESTIONS | - 55 - |
| APPENDIX I: WRITTEN SURVEY | - 57 - |
| APPENDIX J: PARTICIPANT BACKGROUND | - 59 - |
| LIDAR KNOWLEDGE..... | - 59 - |
| APPENDIX K: MEASUREMENT RESULTS (RQ1) | - 61 - |
| MEASUREMENT ESTIMATES PER PARTICIPANT – SCENE A (2D & 3D)..... | - 61 - |
| MEASUREMENT ESTIMATES PER PARTICIPANT – SCENE B (2D & 3D)..... | - 64 - |
| MEASUREMENT COMMENTS | - 67 - |
| FURTHER MEASUREMENT ACCURACY COMMENTS (Q8)..... | - 76 - |
| APPENDIX L: INTERPRETATION RESULTS (RQ2) | - 78 - |
| SUMMARY OF INTERPRETATION ACCURACY RESULTS..... | - 78 - |

| | |
|--|---------------|
| 2D & 3D INTERPRETATION FOR EACH POI IN SCENE C..... | - 79 - |
| 2D & 3D INTERPRETATION FOR EACH POI IN SCENE D..... | - 84 - |
| EXAMPLES OF 2D VS. 3D VERBAL INTERPRETATIONS | - 89 - |
| FURTHER COMMENTS ON INTERPRETATION ACCURACY..... | - 90 - |
| APPENDIX M: REFLECTION ON METHODS (RQ3) | - 91 - |
| GENERAL 2D AND 3D OBSERVATIONS | - 91 - |
| INTERPRETATION FEEDBACK - NATURAL/MAN-MADE FEATURES..... | - 92 - |
| APPENDIX N: DIGITAL APPENDICES | - 95 - |
| README FILE | - 95 - |
| .TXT FILES - AOIS FOR SCENES A, B, C, AND D..... | - 95 - |
| .PY FILES - VISUALISATION MODULES..... | - 95 - |

Appendix A: Lidar data processing

Airborne instrument specification

Optech ALTM Gemini summary specification (Optech, 2008)

Laser Wavelength 1064 nanometers

Range Capture Up to 4, including 1st, 2nd, 3rd, last returns

Intensity Capture 12-bit dynamic range for all recorded returns, including last

Pulse Rate Frequency 33 - 167 kHz

Beam Divergence nominal (full angle) Dual Divergence 0.25 mrad (1/e) or 0.80 mrad (1/e)

Scan FOV 0 - 50°; Programmable in $\pm 1^\circ$ increments

Scan Frequency 0 – 70Hz (>70Hz optional). Programmable in 1Hz increments

Laser Classification Class IV (FDA 21 CFR)

Horizontal Accuracy 1/5,500 x altitude (m AGL); 1 sigma

Elevation Accuracy 5 - 35 cm; 1 sigma

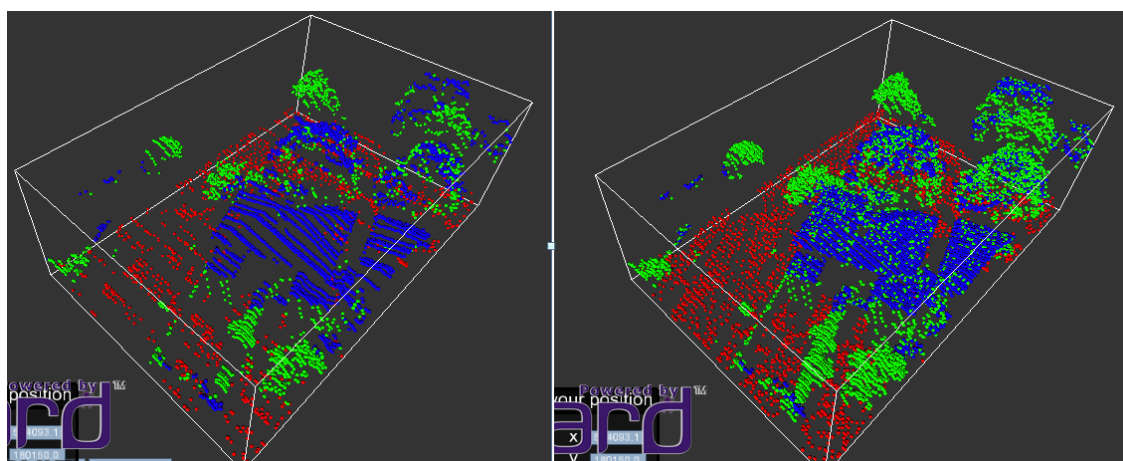
Scanner Product Up to Scan angle x Scan frequency = 1000

Roll Compensation $\pm 5^\circ$ at full FOV – more compensation available if FOV reduced

Position Orientation System POS/AV 510 OEM includes embedded BD960 GNSS receiver (GPS and GLONASS)

Operational Altitudes 150-4000 m, Nominal

Screenshot of London lidar data issue



When visualised, Scene C exhibited gaps in the data (left), so the data were reprocessed using returns 1, 2, and 3 (result shown in right-hand image). Lidar data © Airbus Defence and Space Ltd. (2013a). Vizard 3.0 visualisation software © Worldviz (2010).

LAStools script

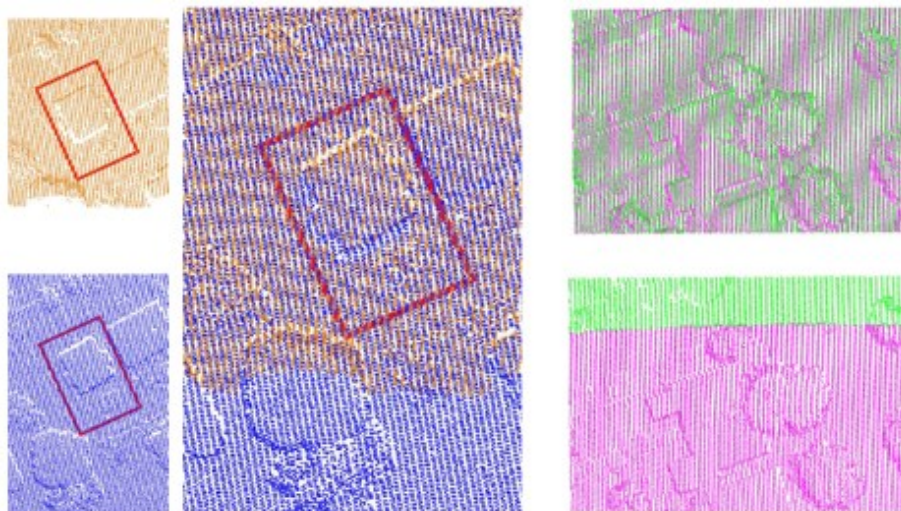
Processing steps for the creation of Scenes A, B, C, and D using lasclip and las2txt open-source processing applications from LAStools lidar software (Isenburg, 2013).

| Process Step | A | B | C | D |
|---|---|---|---|--|
| File Input (lidar data file, .las) | ST575725_edited_RGB | ST575725_edited_RGB | TQ240800_c1 23456_15normal | ST565725_classified_RGB |
| Code to clip .las to AOI (.shp or min and max xy) using LAStools (lasclip or las2las functions) | lasclip -i ST575725_edited_RGB.las -poly house_polygon.shp -verbose | las2las -i ST575725_edited_RGB.las -o ST575725_edited_tree.las -clip 357840 172582 357884 172620 | las2las -i TQ240800_c1 23456_15normal.las -o TQ240800_c1 23456_15clip.las -clip 524108.24 180159.90 524161.14 180194.40 | lasclip -i ST565725_classified_RGB.las -poly ST565725cliff.shp -verbose |
| Code to convert clipped .las file to .txt format using LAStools (las2txt function) | las2txt -i ST575725_edited_RGB_1.las -o BEd26SceneA.txt -parse xyziRGBc -drop_class 1 5 -sep comma | las2txt -i ST575725_edited_tree.las -o BEd25SceneB.txt -parse xyziRGBc -drop_class 1 6 -sep comma | las2txt -i TQ240800_c1 23456_15clip.las -o LHT1thc5615d rop1234.txt -parse xyziRGBc -drop_class 1 2 3 4 -sep comma | las2txt -i ST565725_classified_RGB_1.las -o BCLF1thrgbcN O1.txt -parse xyziRGBc -drop_class 1 -sep comma |
| Pseudocode to convert clipped .las file to .txt format using LAStools las2txt* | write to .txt as x,y,z,i,R,G,B,c but omit unclassified (1) and high vegetation (5) points | write to .txt as x,y,z,i,R,G,B,c but omit unclassified (1) and building points (6) | write to .txt as x,y,z,i,R,G,B,c but omit classes 1, 2, 3, 4 [returns 1-4] | write to .txt as x,y,z,i,R,G,B,c but omit unclassified points (1) |
| File output (AOI of point cloud, .txt) | BEd26SceneA | BEd25SceneB | LHT1thc5615d rop1234 | BCLF1thrgbcN O1 |
| Number of points in file* | 2,305 | 4,396 | 10,628 | 14,574 |
| Number of points in each classification | 824 Ground 1323 Building | 1699 Ground 2666 Veg | 1563 Ground 2450 Veg 2991 Building | 6209 Ground 7242 Veg 1123 Building |

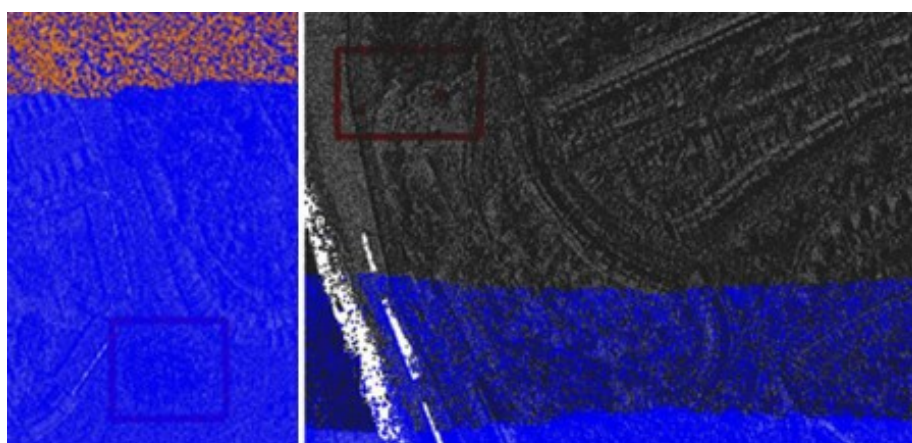
**Files contain other classifications that are not detailed here*

Flightline screenshots

Overlap between Flightlines (left: Scene A, right: Scene C)

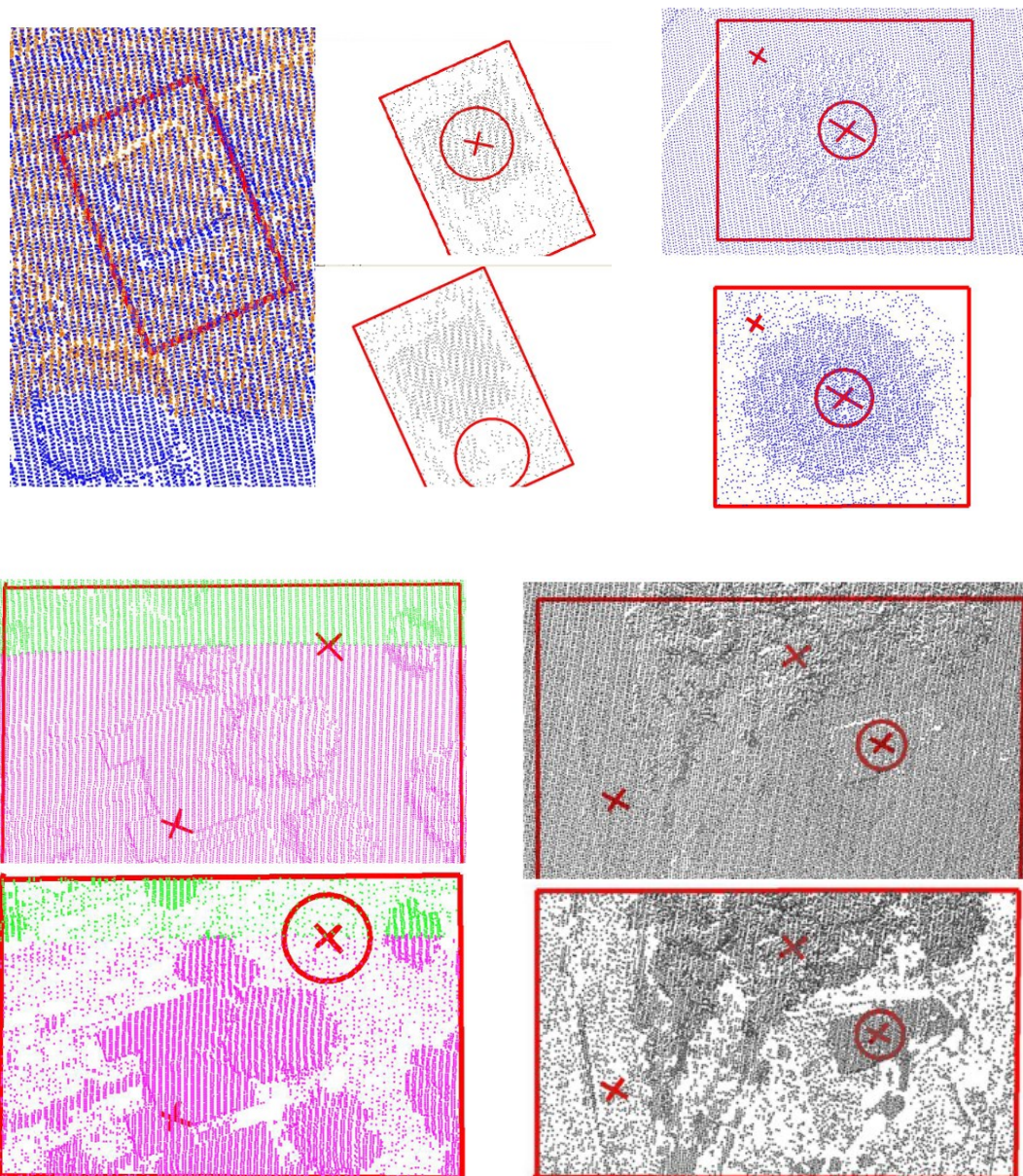


No overlap between Flightlines (left: Scene B, right: Scene D)



Screenshots of airborne lidar survey flightlines. Top row from left: Scene A, made up for two overlapping flightlines. Scene C was originally made up for 2 flightlines, but the overlapping points were filtered. Bottom row from left: Scene B (left) and D (right) coloured by flightline – no overlap. Flightlines viewed using Terrasolid (2013). Lidar data © Airbus Defence and Space Ltd. (2013a & 2013b).

Density screenshots



Screenshots of raw lidar coverages for each AOI vs. processed coverage. Each circle is 10m in diameter indicates location of density sample areas, crosses denote centre of sample areas. **Top row from left:** Scene A, main image, showing raw data coloured by overlapping flightlines. Subsets show processed AOI made up of building and ground points. Circles indicate areas over building (top) and ground (bottom) where point density was reported; Scene B shown with raw data (top) and classified data (bottom) made up of vegetation and ground points. **Bottom row from left:** Scene C coloured by flightline in both unprocessed (top) and processed (bottom) AOI; Scene D shown unprocessed (top) and processed (bottom). Flightlines viewed using Terrasolid (2013). Lidar data © Airbus Defence and Space Ltd. (2013a & 2013b).

Appendix B: Visualisation code

List of main modules

List of all the main Python (.py) modules used for 2D and 3D visualisations.

| | Task: | Measurement | | Interpretation | |
|--------|--------|--------------|--------------|----------------|--------------|
| | Scene: | A | B | C | D |
| Method | 2D | SceneA_2D_00 | SceneB_2D_00 | SceneC_2D_01 | SceneD_2D_01 |
| | 3D | SceneA_3D_00 | SceneB_3D_00 | SceneC_3D_01 | SceneD_3D_01 |

List of modules called into main modules

| Module Name | Function | Imports | File Output |
|-----------------------|--|-------------------|----------------------------|
| MainViewModVaryd | Sets aspects of Mainview on keydown 2,4,8,6 | xyzrangesmins.txt | n/a |
| VerticesMod | Create a bounding box around the point cloud. On key down 'c', create cube On key down 'g' create grid | xyzrangesmins.txt | n/a |
| GamepadMod | Updates according to gamepad interaction, appending button press and time to measurelog. | xyzrangesmins.txt | measurelog.txt |
| MouseStateMod | states mouse use (of interviewer) | | |
| KeyStateMod | states keydown (of interviewer) | | |
| TargetXYZtoFileMod | Output selected points (pointcoordstr1 and pointcoordstr2) to file | n/a | measurelog.txt |
| UserPosiMod | | xyzrangesmins.txt | n/a |
| UserTrackerMod<scene> | Track user position. Get mainview position and euler, write to file with timestamp and participant no. | ParticipantNo.txt | tracking_+str(subject).txt |

Main module (.py)

Example

```

#-----
# FILENAME: SceneA_2D_00 (MainMod)
# CB
# Updated 131027
#-----

#-----
viz.setMultiSample(4) # Fullscreen anti-aliasing samples to perform
after each frame is rendered.
#-----

#=====
#-----
viz.go() #2D METHOD

viz.fov(60,1.29435) #viz.fov(40,1.29435) # true width = 4m
#<viz>.fov( verticalFOV [vert. FOV in degrees], aspectRatio [horz. to
vert. aspect ratio])

# VR screen vertical fov = 40.0, VR screen horizontal fov =
51.7740087592, aspectRatio = 51.77/40.0
#-----

#-----
#viz.go(viz.QUAD_BUFFER) #3D METHOD
#-----
#=====

# Import built-in Vizard Modules
import viz, vizact, vizcam, vizinfo, vizmenu, vizshape, viztask, math,
vizjoy

#Import bespoke mods
#import View3DMod # <-3D ONLY (sets IPD and FOV)
import UserPosiMod
import GamepadMod
import LightMod

joy = vizjoy.add() #Add a joystick

# declare static variables
global normalzview
normalzview = 1.82 # standard height set for user

#-----outputlog.txt set-up -----
#Declare
global outputlog
outputlog = open('outputlog.txt','a') #a for append.

participantfile = open('ParticipantNo.txt','r')
for i in participantfile:

```

```

    subject = i # input the participants' number
    print i

import datetime
time = datetime.datetime.now()
timestamp = str(time)
scenetimestamp = '\n' + timestamp + ', ' + str(subject) + ',
SceneA_2D_00' + '\n'
print scenetimestamp
outputlog.write(scenetimestamp)

#store the time at which this trial started
startTime = viz.tick()
#-----outputlog.txt set-up -----

def guisettings(): # user screen characteristics

    #Set window name
    viz.window.setName('Scene A, 2D')

    #Set the background colour
    colormax = 255
    bkcolr = 178/float(colormax) #light green r
    bkcolg = 223/float(colormax) #light green g
    bkcolb = 138/float(colormax) #light green b
    viz.clearcolor([bkcolr,bkcolg,bkcolb]) # colorbrewer, 4 colours,
quant,

    return

guisettings()

#-----
def lidarin(): #read in lidar data file and allocate its content to
arrays

    # Open the relevant data file
    lidarfile = open('BEd26SceneA.txt','r') # Bristol, just classes 2,6

    #Define arrays to be used to allocate values from file
    x = []          # defines empty array for x
    y = []          # y
    z = []          # z
    i = []          # intensity
    r = []          # r
    g = []          # g
    b = []          # b
    c = []          # c
    rn = []        # return number

    global sumlines
    sumlines = 0 # define the counter for number of lines

    global adjusted_x
    global adjusted_y
    global adjusted_z
    adjusted_x = []
    adjusted_y = []
    adjusted_z = []

```

```

#Set initial bounding values

max_x = float(-100000000)
min_x = float(100000000)

max_y = float(-100000000)
min_y = float(100000000)

max_z = float(-100000000)
min_z = float(100000000)

max_i = float(-100000000) #if intensity
min_i = float(100000000)

max_r = float(-1) #if rgb
min_r = float(256)

max_g = float(-1) #if rgb
min_g = float(256)

max_b = float(-1) #if rgb
min_b = float(256)

#Loop round the file, line by line
for line in lidarfile:

    #import CoordAdjustMod

    print '-----'

    # true_ the line into component parts
    #s = line.split( ) # if spaces
    s = line.split(',') #if comma delimited

    #Allocate columns to true_x etc
    true_x = (float(s[0])) #int(s) converts the input str into
integer
    x.append(true_x)

    if true_x < min_x:
        min_x = true_x
    if true_x > max_x:
        max_x = true_x

    print 'true_x=', true_x
    print 'min_x=', min_x
    print 'max_x=', max_x
    #print 'coordsx = ', coordsx
    print 'adjusted_x = ', adjusted_x
    print 'done max min x for line', sumlines

    #===== Remembering
    # Vizard's y coordinate is really 'z' in the real world,
so allocate the true zs to y array.
    # i.e. true xyzi = Vizard xzyi, so [0 = x, 2 = y, 1 =
z, 3 = i]
    #=====

```



```

true_y = (float(s[1]))
y.append(true_y)

if true_y < min_y:
    min_y = true_y

if true_y > max_y:
    max_y = true_y

print 'true_y=', true_y
print 'min_y=', min_y
print 'max_y=', max_y
print 'adjusted_y = ', adjusted_y
print 'done max min y for line',sumlines

true_z = (float(s[2]))
z.append(true_z)

if true_z < min_z:
    min_z = true_z

if true_z > max_z:
    max_z = true_z

print 'done max min z for line',sumlines

true_i = float(s[3]) #if intensity values are in the file:
i.append(true_i)

if true_i < min_i:
    min_i = true_i
if true_i > max_i:
    max_i = true_i

print '(done max min i for line',sumlines,')'

true_r = float(s[4]) #if rgb values available
r.append(true_r)

if true_r < min_r:
    min_r = true_r
if true_r > max_r:
    max_r = true_r

print '(done max min r for line',sumlines,')'

true_g = float(s[5]) #if rgb values are in the file:
g.append(true_g)

if true_g < min_g:
    min_g = true_g
if true_g > max_g:
    max_g = true_g

print '(done max min g for line',sumlines,')'

true_b = float(s[6]) #if rgb values are in the file:
b.append(true_b)

```

```

        if true_b < min_b:
            min_b = true_b
        if true_b > max_b:
            max_b = true_b

        print '(done max min b for line',sumlines,')'

        true_c = float(s[7]) #if class values are in the file:
        c.append(true_c)
#         true_rn = float(s[8]) #if return values are in the file:
#         rn.append(true_rn)

        print '(read class for line',sumlines,')'

        print 'xyz(i/rgb) values for line number', sumlines, ':',
line
        sumlines += 1

#Declare variables for range values
global x_range, y_range, z_range, i_range
global r_range, g_range, b_range

x_range = max_x - min_x
y_range = max_y - min_y
z_range = max_z - min_z
i_range = max_i - min_i
r_range = max_r - min_r
g_range = max_g - min_g
b_range = max_b - min_b

#=====
# Create xyzrangesmins.txt : # Write xy+z minimum values and ranges
to file so that can be used
#for the bounding box grid construction (VerticesMod) and mainview
positioning (MainViewMod)

# Opens file 'response.txt' in write mode
file = open('xyzrangesmins.txt', 'w')

# Create the output string as follows : 0-2 = minxyz; 3-5 = ranges
of xyz
out = str(min_x) + ',' + str(min_y) + ',' + str(min_z) + ',' +
str(x_range) + ',' + str(y_range) + ',' + str(z_range)

# Write the string to the output file
file.write(out)
#=====

#-----
#Define empty array for the lidar hits
lidar_hits = []

# For hits in file

```

```

for hit in range(sumlines):

    #Create a ball
    #ball = viz.add('red_ball.wr1')
    ballmono = viz.add('red_ball.wr1')
    ballcolor = viz.add('red_ball.wr1')

    #Set its scale
    #ball.scale(1.0,1.0,1.0) #ALS size
    #ballmono.scale(2.0,2.0,2.0) #ALS size DUT
    ballcolor.scale(1.5,1.5,1.5)
    ballmono.scale(1.5,1.5,1.5)

    #-----
    # Set position
    from math import sqrt, pow #math.sqrt(x) #math.pow(x, y)

    global nonnegx, nonnegy, nonnegz

    #true_x value
    nonnegx = (x[hit] - min_x)
    #true_y value
    nonnegy = (y[hit] - min_y)
    #true_z value
    nonnegz = (z[hit] - min_z)

    ballmono.setPosition([nonnegx,nonnegz,nonnegy]) # reflect
Vizard's framework
    ballcolor.setPosition([nonnegx,nonnegz,nonnegy]) # reflect
Vizard's framework
    #-----
    # Set colour of ball

    #Mono
    global mcolr
    global mcolg
    global mcolb

    mcolr = []
    mcolg = []
    mcolb = []

    colormax = 255

#    mcolr = 51/float(colormax) # dark green
#    mcolg = 160/float(colormax) # dark green
#    mcolb = 44/float(colormax) # dark green

    mcolr = 31/float(colormax) #dark blue
    mcolg = 120/float(colormax) # dark blue
    mcolb = 180/float(colormax) # dark blue

    ballmono.color(mcolr,mcolg,mcolb) #130808
    ballmono.visible(viz.ON)

    #ASPRS official las classifications
    #0 Created, never classified
    #1 Unclassified
    #2 Ground

```

```

#3 Low Vegetation
#4 Medium Vegetation
#5 High Vegetation
#6 Building
#7 Low Point (noise)
#8 Model Key-point (mass point)
#9 Water

if c[hit] == 2: #2asprs = Ground

    gcolr = 31/float(colormax) #dark blue
    gcolg = 120/float(colormax) # dark blue
    gcolb = 180/float(colormax) # dark blue
    ballcolor.color([gcolr,gcolg,gcolb]) # 25-10-13
    ballcolor.visible(viz.OFF)

# elif c[hit] == 5: #5 asprs, = High Veg
#     vcolr = 51/float(colormax) # dark green
#     vcolg = 160/float(colormax) # dark green
#     vcolb = 44/float(colormax) # dark green
#     ballcolor.color([vcolr, vcolg, vcolb]) # 25-10-13
#     ballcolor.visible(viz.OFF)

elif c[hit] == 6: # 6asprs = Building
    bcolr = 116/float(colormax) # light blue
    bcolg = 206/float(colormax) # light blue
    bcolb = 227/float(colormax) # # light blue
    ballcolor.color([bcolr, bcolg, bcolb]) # 25-10-13
    ballcolor.visible(viz.OFF)

#Append ball object to array
#lidar_hits.append(ball)
lidar_hits.append(ballmono)
lidar_hits.append(ballcolor)

lidarfile.close() # the stimulus file is closed

return

lidarin()

#-----

#===Import more bespoke modules=====
import VerticesMod # for bounding box 'c' & grid 'g'
import UserTrackerMod_sceneA_2D # to track the user position
import TargetXYZtoFileMod # for measuring
import MouseStateMod # states mouse use
import KeyStateMod # states keydown
#===Import more bespoke modules=====

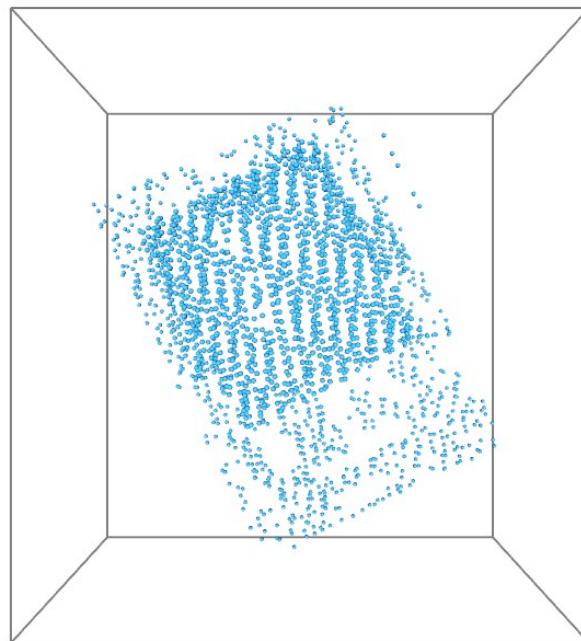
outputlog.close

```

Appendix C: Visualised datasets

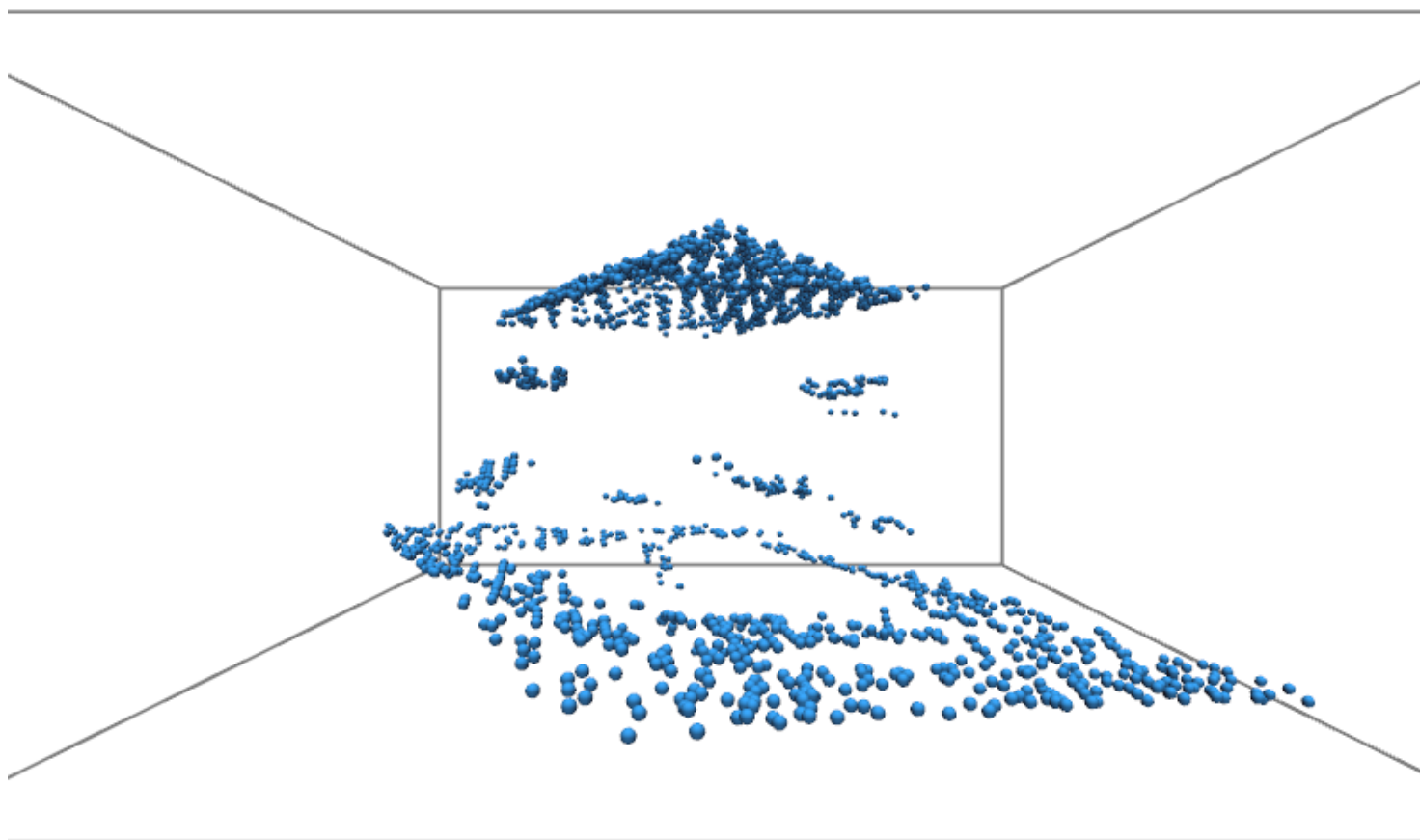
Scene A – point cloud dataset

Plan View



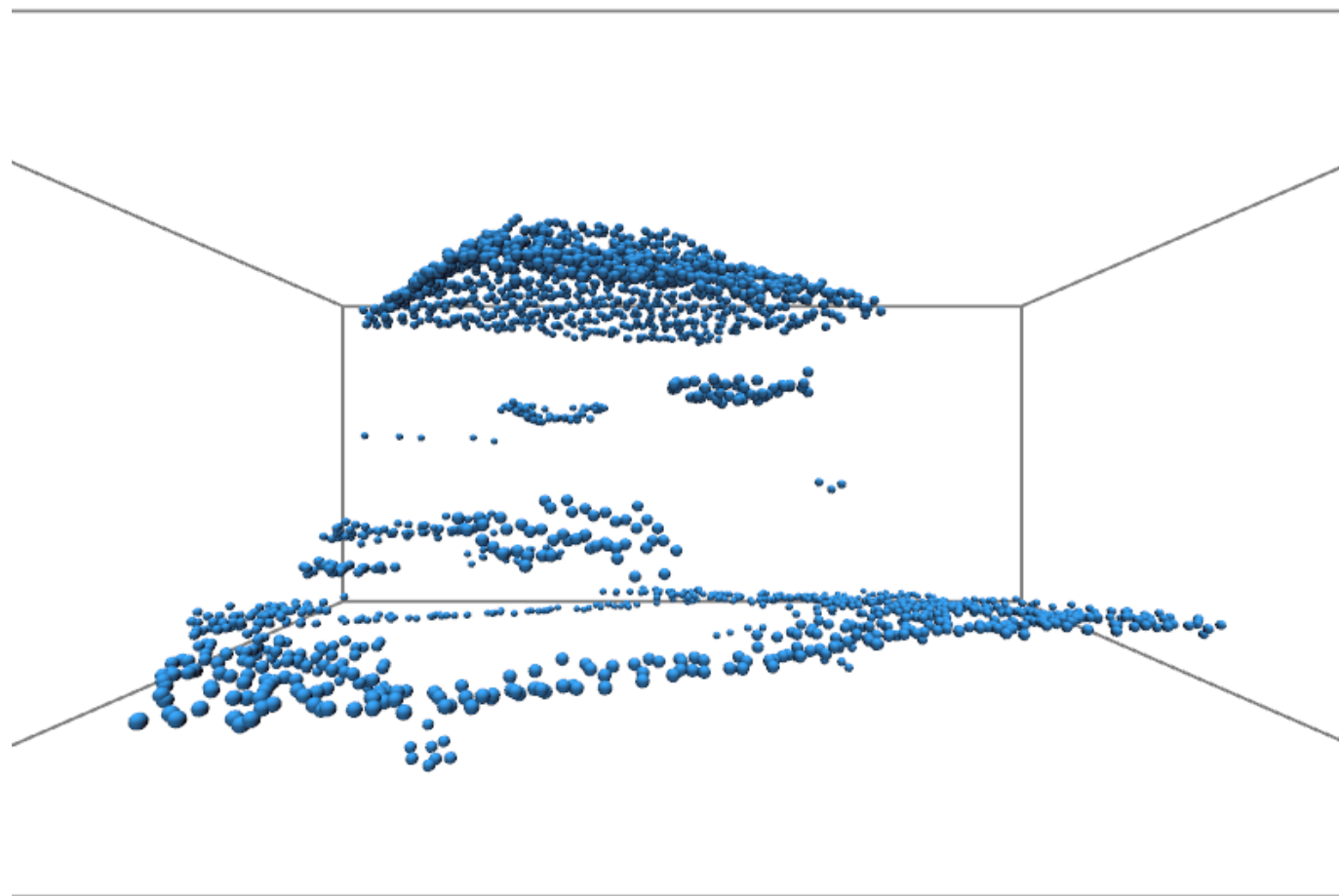
Plan view of Scene A, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (30m x 34m x 15m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene A – North-facing



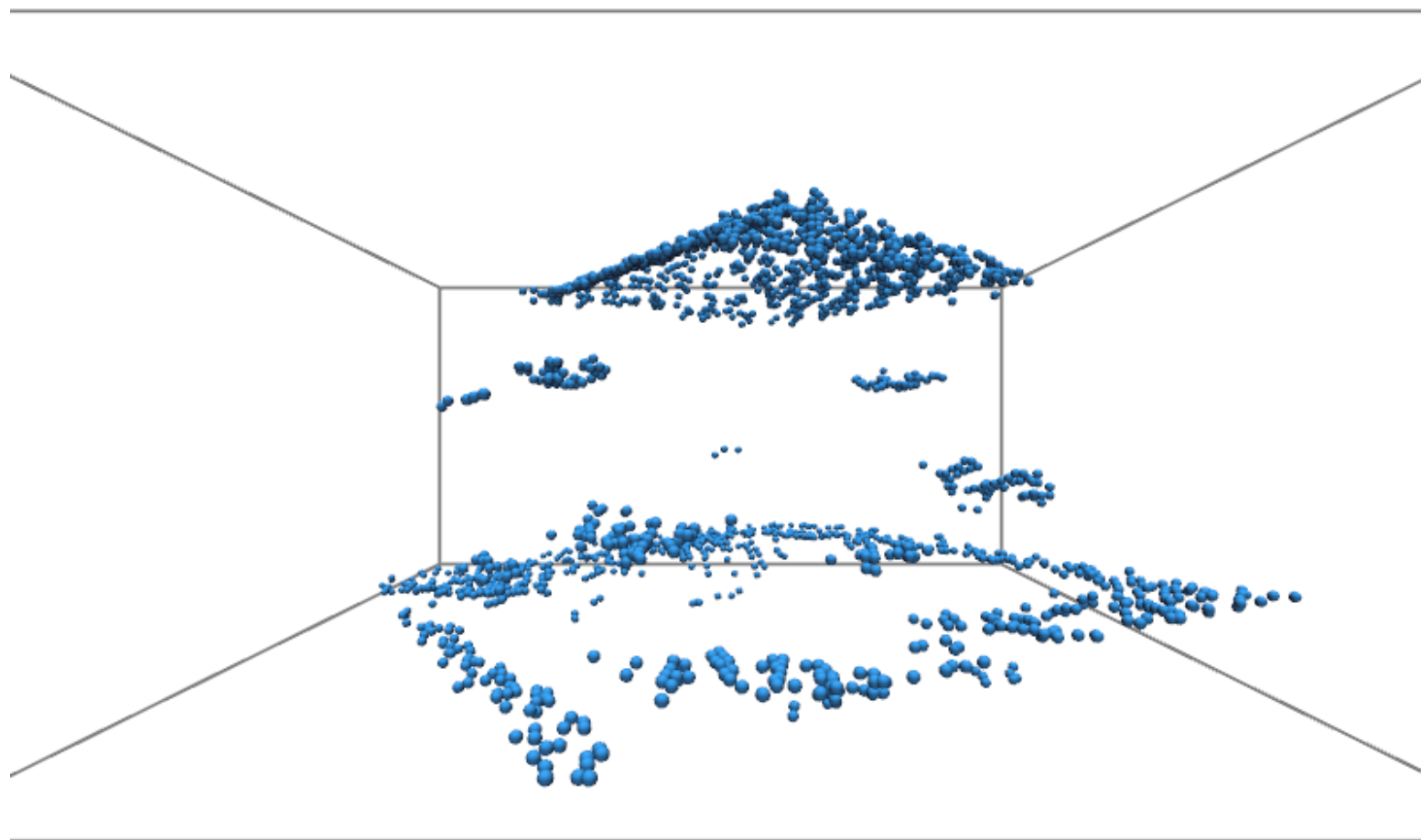
North-facing view of Scene A, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (30m x 34m x 15m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene A – East-facing



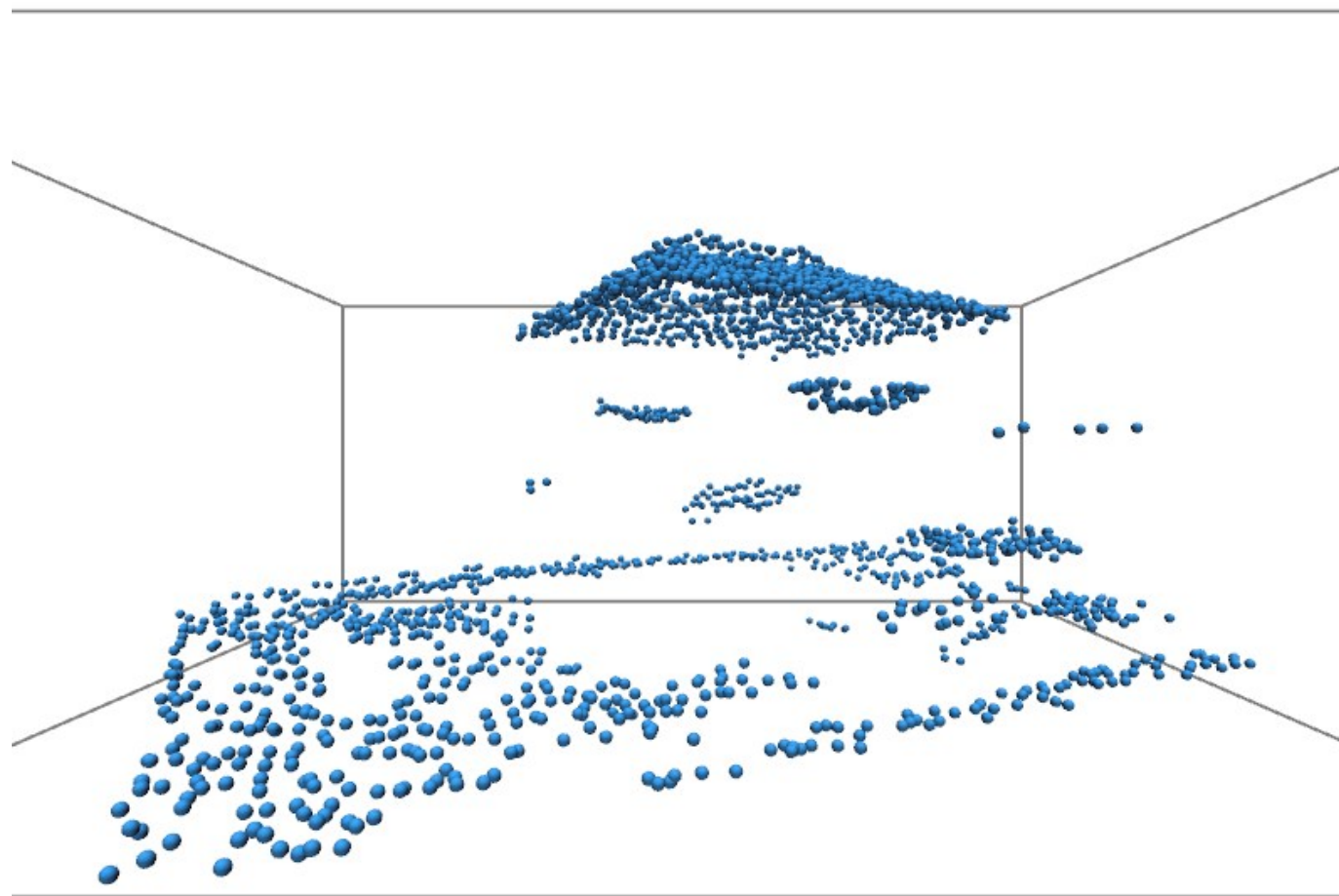
East-facing view of Scene A, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (30m x 34m x 15m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene A – South-facing



South-facing view of Scene A, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (30m x 34m x 15m). Lidar data © Airbus Defence and Space Ltd. (2013a).

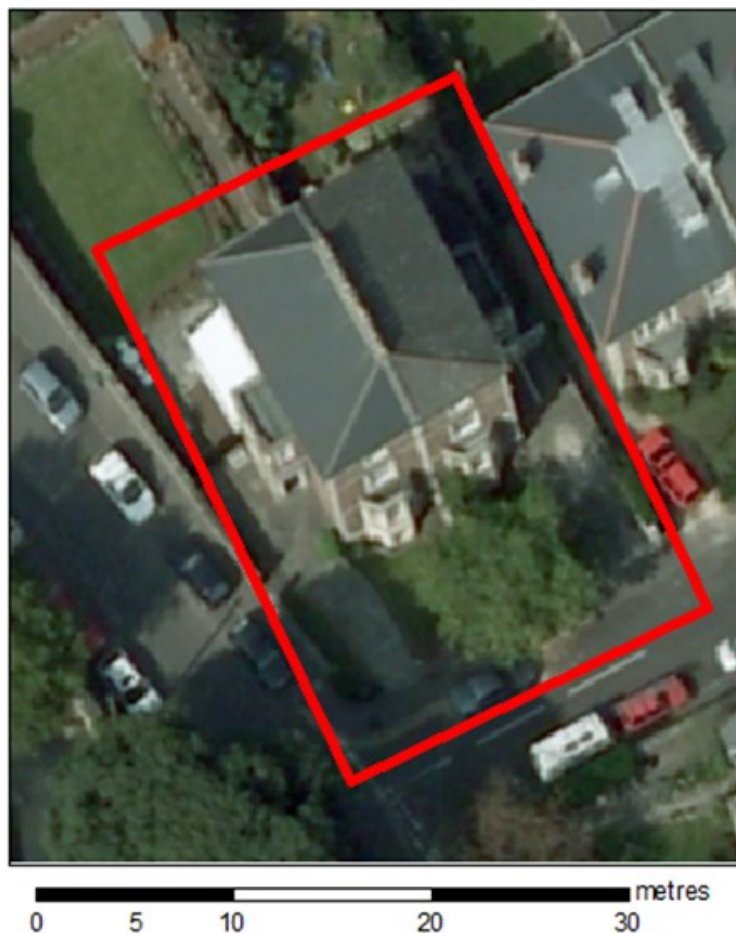
Scene A – West-facing



West-facing view of Scene A, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (30m x 34m x 15m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene A – Reference photos shown to participants

Orthophotograph



Aerial photography of Scene A, taken 14 April 2007, during a survey of the Bristol area (OS tile ST5672). This image was shown to participants after completion of the measurement task, to help place the point cloud in its real-world context. Orthophotograph © GeoPerspectives (2013a).

StreetView photo 1



StreetView imagery © Google (2014), acquired 2012, of Scene A building feature. Image was shown to participants after completion of measurement task. <https://www.google.co.uk/maps/@51.451896,-2.6120702,3a,75y,9.4h,87.12t/data=!3m7!1e1!3m5!1s3qvsQBqBB10WOC4hEa-qYQ!2e0!5s20120501T000000!7i13312!8i6656>

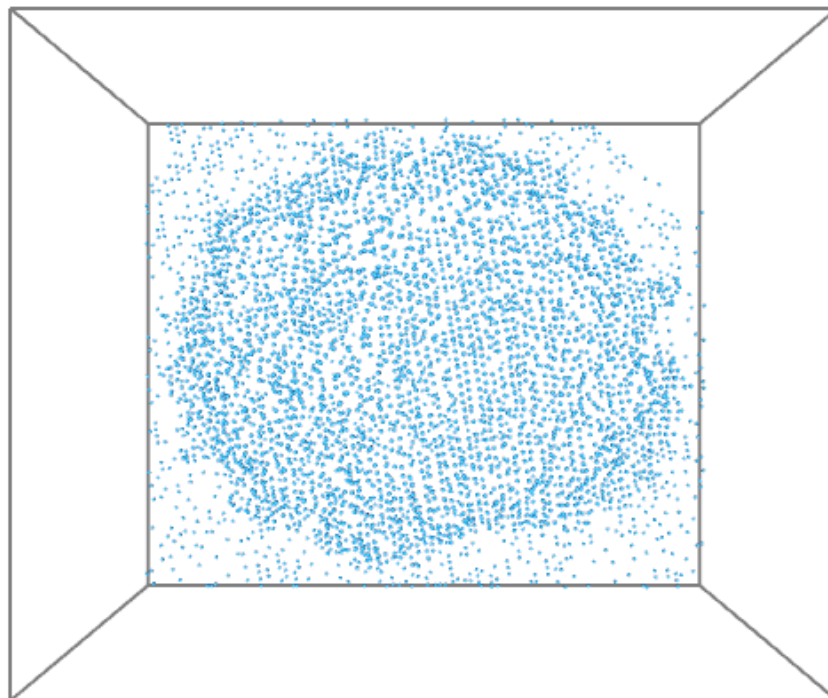
StreetView photo 2



StreetView imagery © Google (2014), acquired 2012, of Scene A building feature. <https://www.google.co.uk/maps/@51.4519464,-2.6122162,3a,75y,29.11h,99.25t/data=!3m7!1e1!3m5!1sWAP6t1F34PSDvjoGu4NHRA!2e0!5s20120501T000000!7i13312!8i6656>

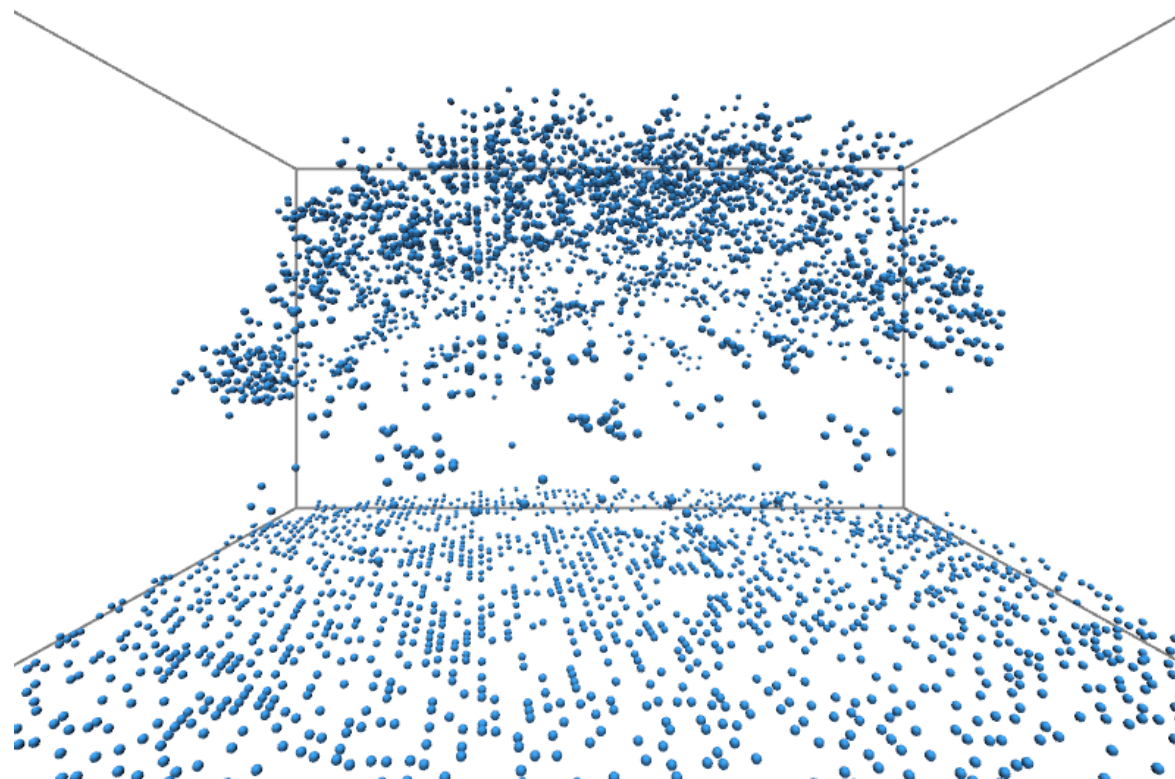
Scene B– point cloud dataset

Plan View



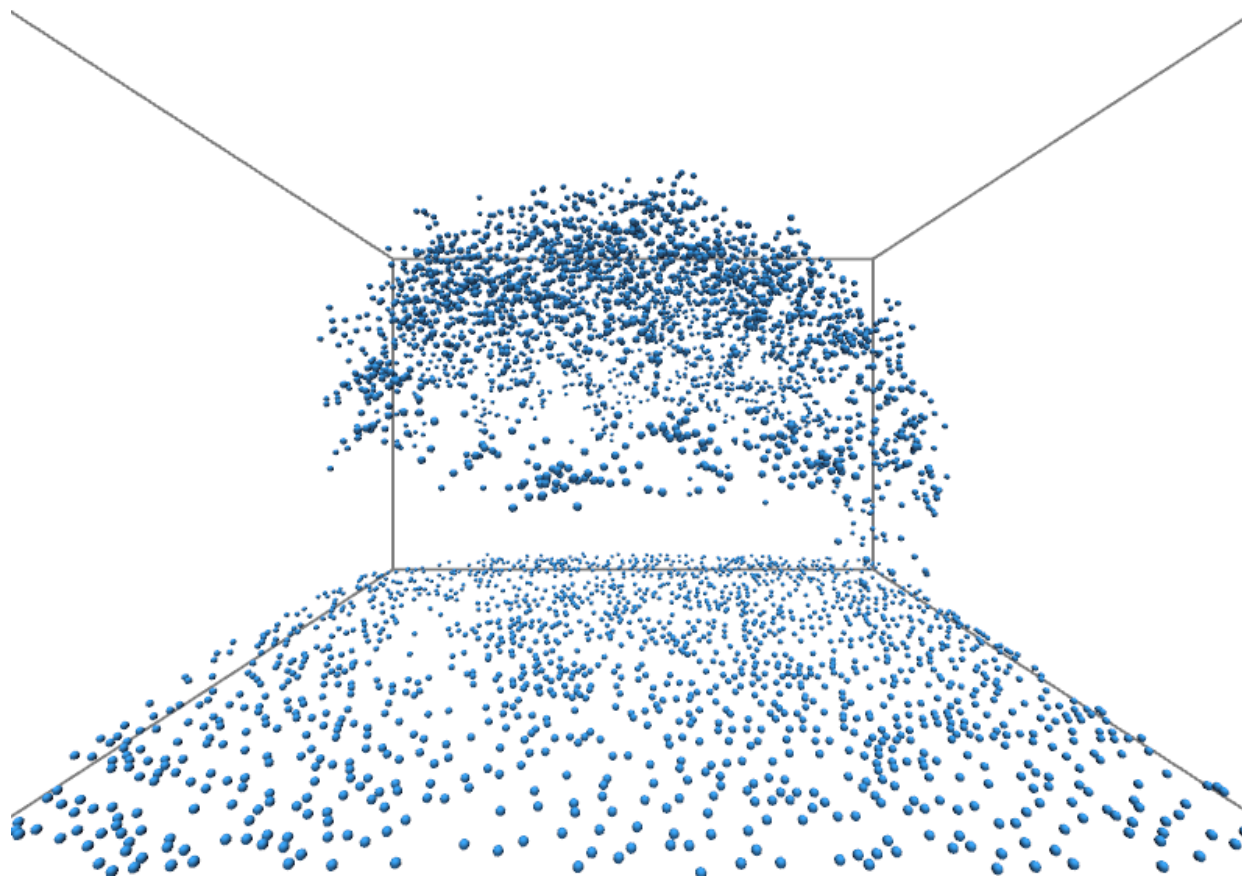
Plan view of Scene B, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (44m x 38m x 25m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene B – North-facing



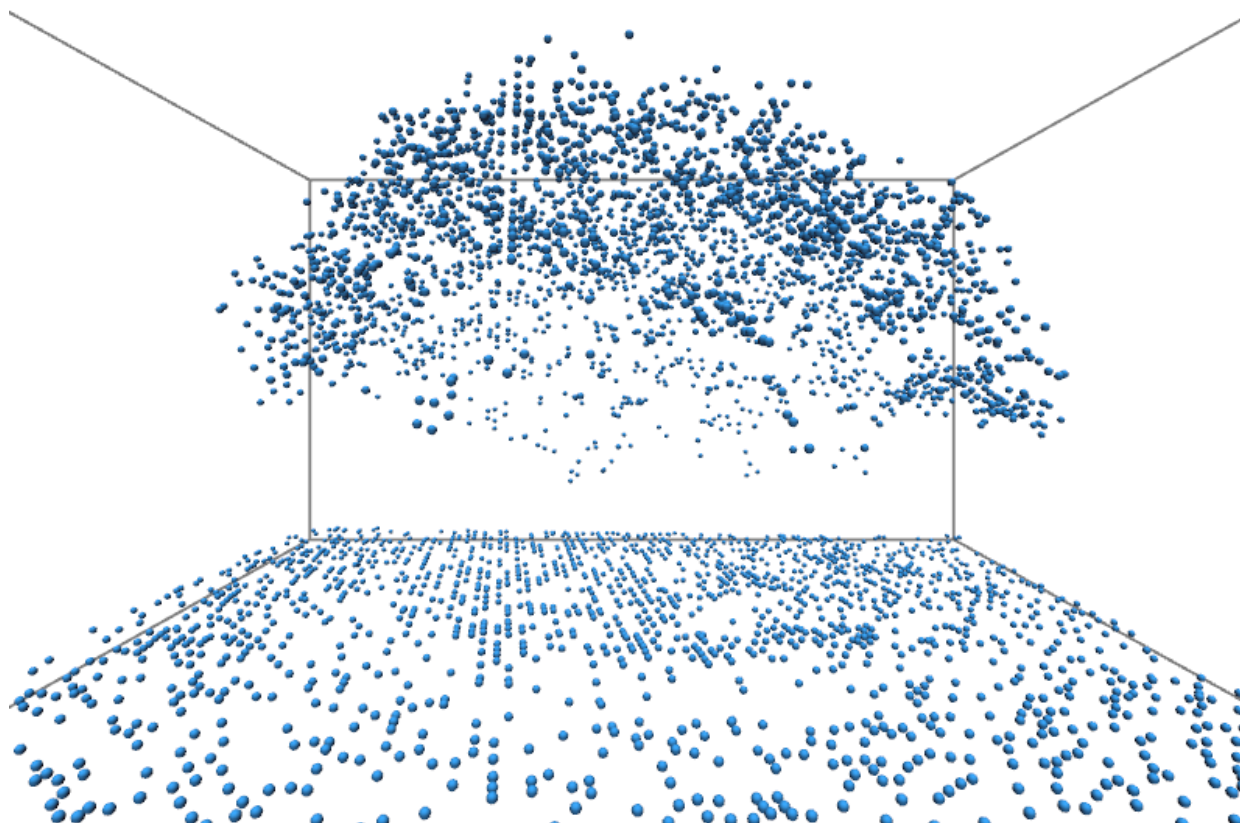
North-facing view of Scene B, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (44m x 38m x 25m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene B – East-facing



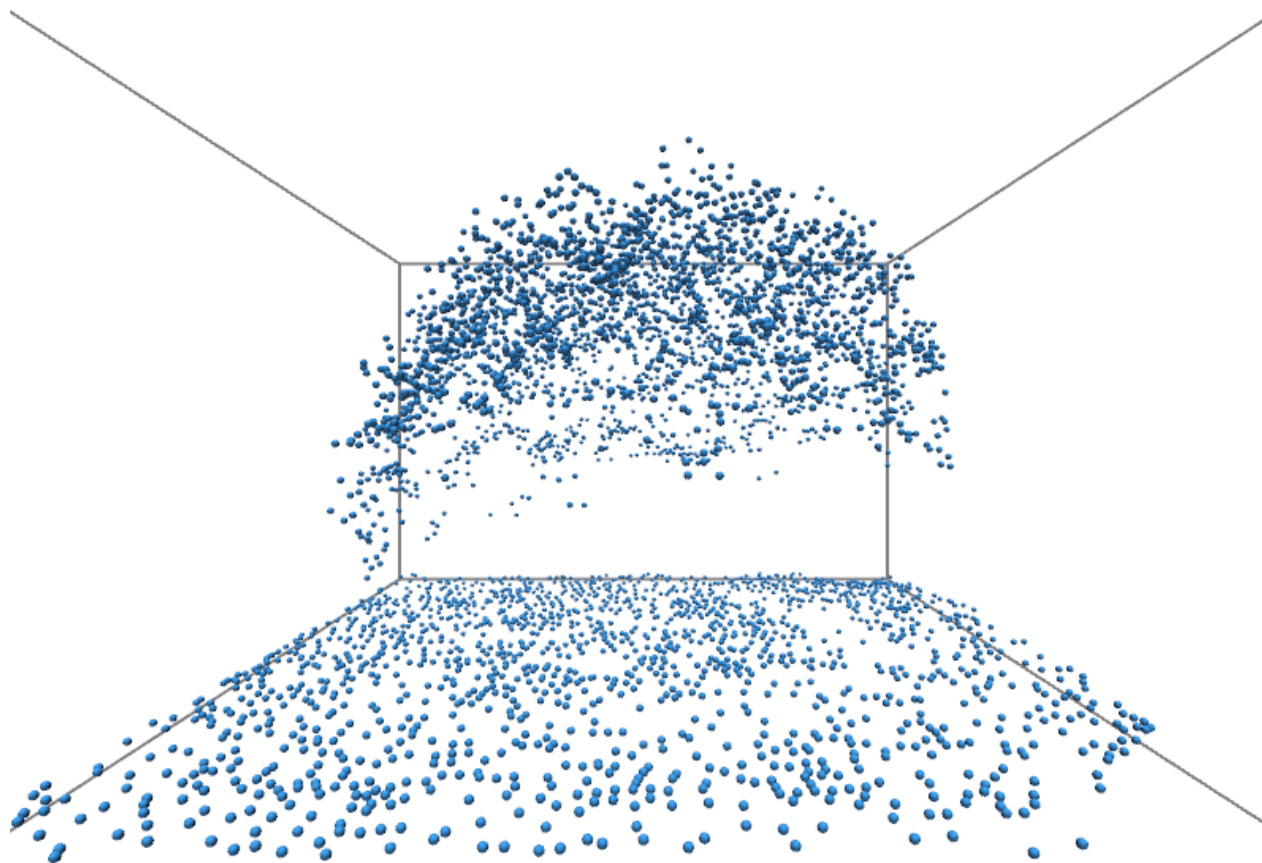
East-facing view of Scene B, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (44m x 38m x 25m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene B – South-facing



South-facing view of Scene B, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (44m x 38m x 25m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene B – West-facing



West-facing view of Scene B, which was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (44m x 38m x 25m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene B – Reference photos shown to participants

Orthophotograph



Aerial photography of Scene B, captured 14 April 2007, during a survey of the Bristol area (OS tile ST5672). Canopy feature is situated on a roundabout. Orthophotograph © GeoPerspectives (2013a).

Scene B - StreetView photo 1



StreetView imagery © Google (2014), captured May 2012, of Scene B vegetation. Participants were often surprised to see several trees. <https://www.google.co.uk/maps/@51.4507459,-2.6073968,3a,75y,281.07h,107.78t/data=!3m7!1e1!3m5!1sJYPrUZl2nY4e0HiCJXHw!2e0!5s20120501T00000!7i13312!8i6656?hl=en-GB>

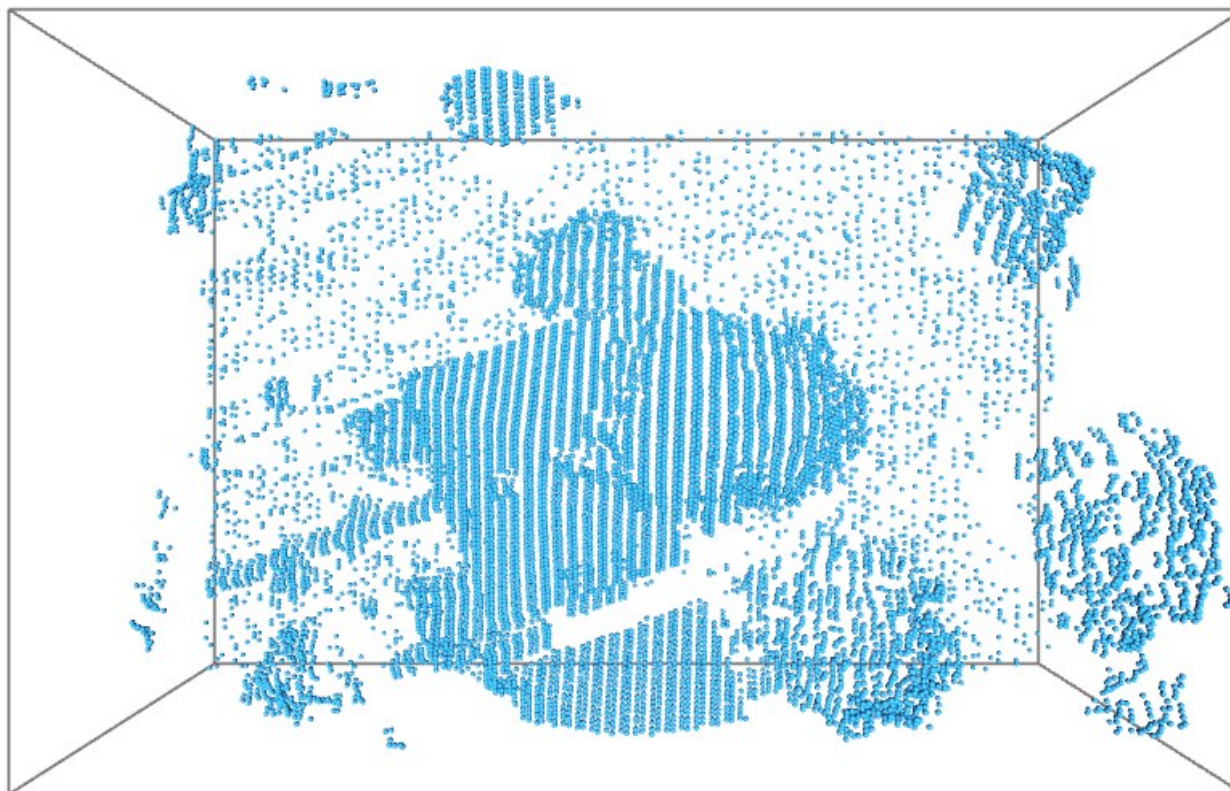
Scene B - StreetView photo 2



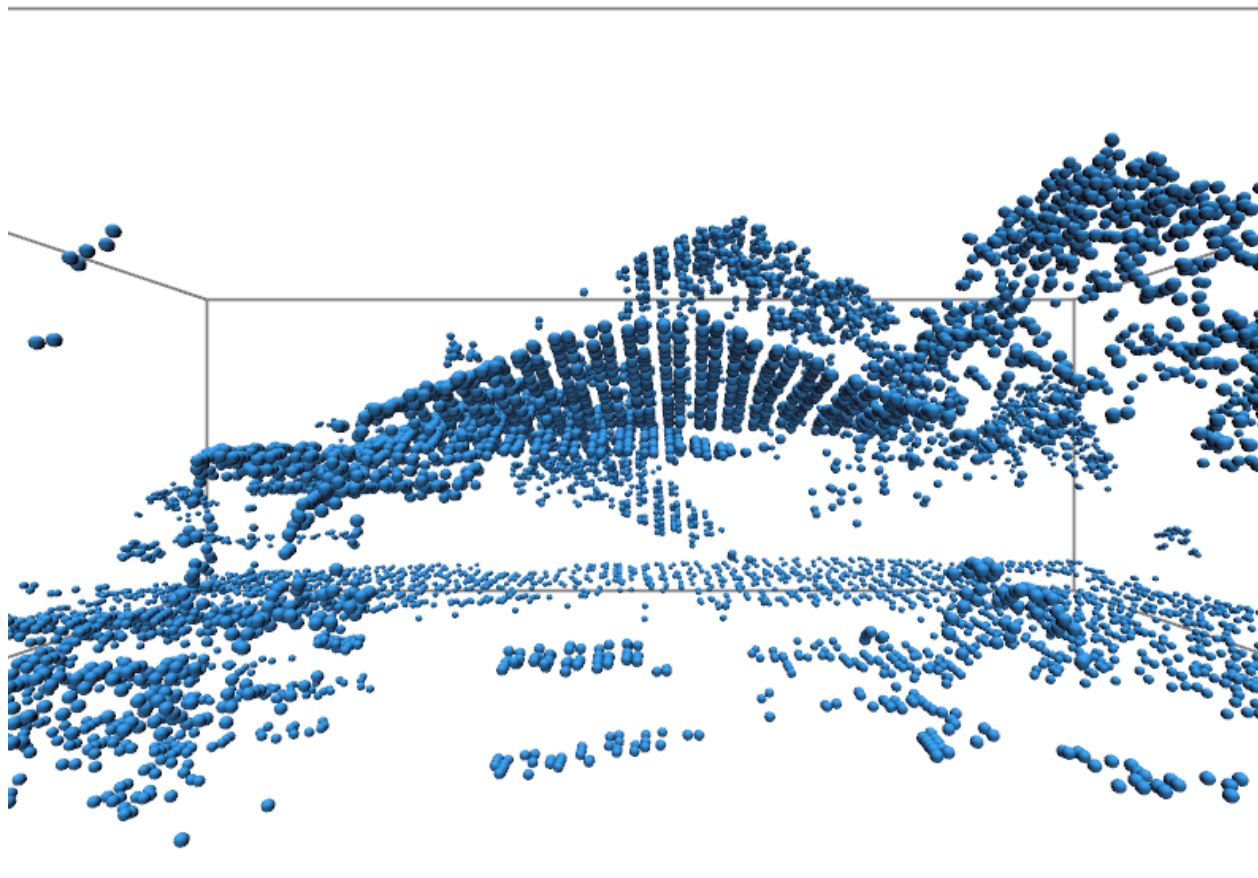
StreetView imagery © Google (2014), captured 2012, showing Scene B vegetation. <https://www.google.co.uk/maps/@51.4510533,-2.6077761,3a,75y,180.96h,102.23t/data=!3m7!1e1!3m5!1sHXftLyG0J5xsd4HnOunm9w!2e0!5s20120501T000000!7i13312!8i6656?hl=en-GB>

Scene C – point cloud dataset

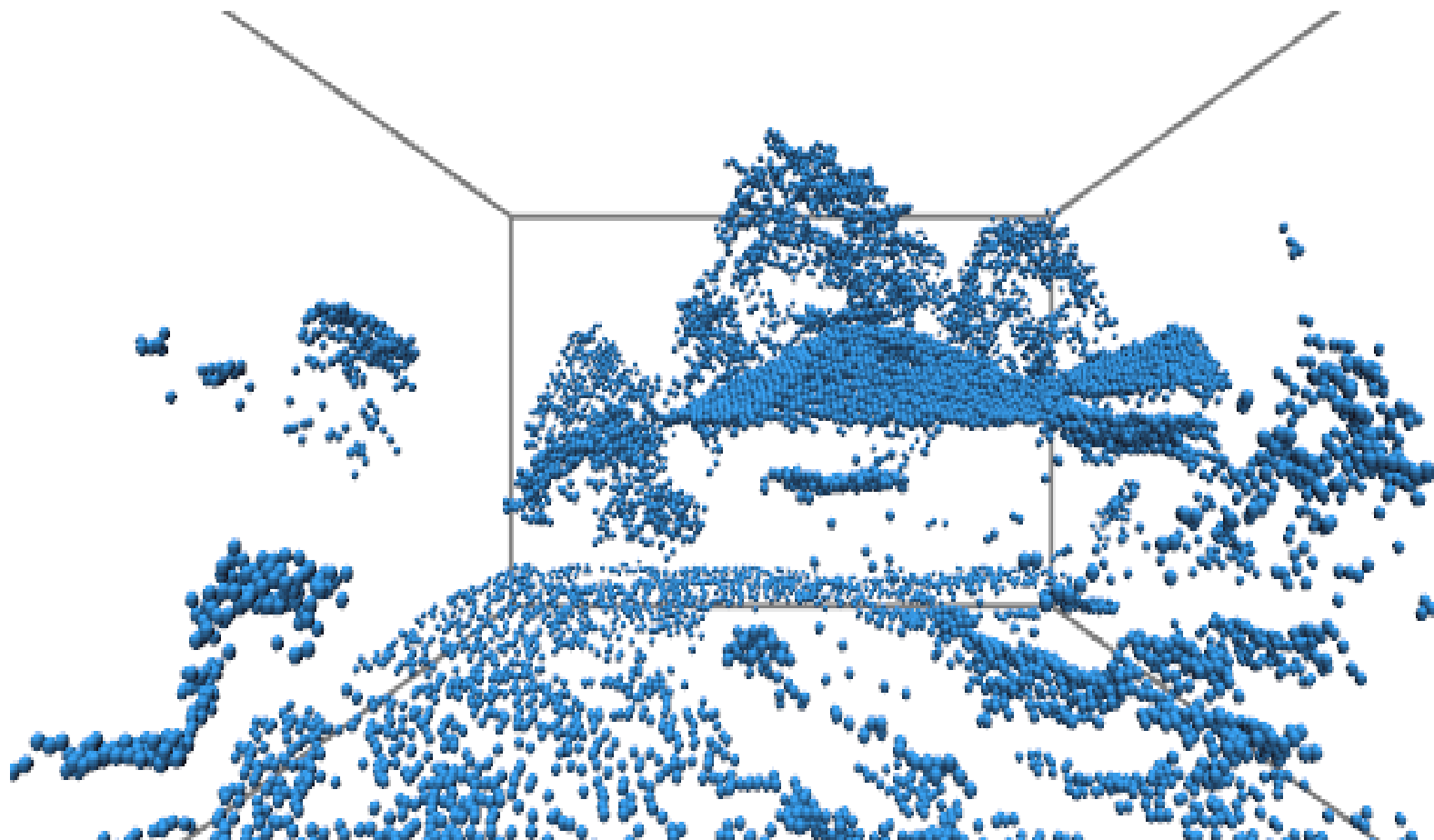
Plan View



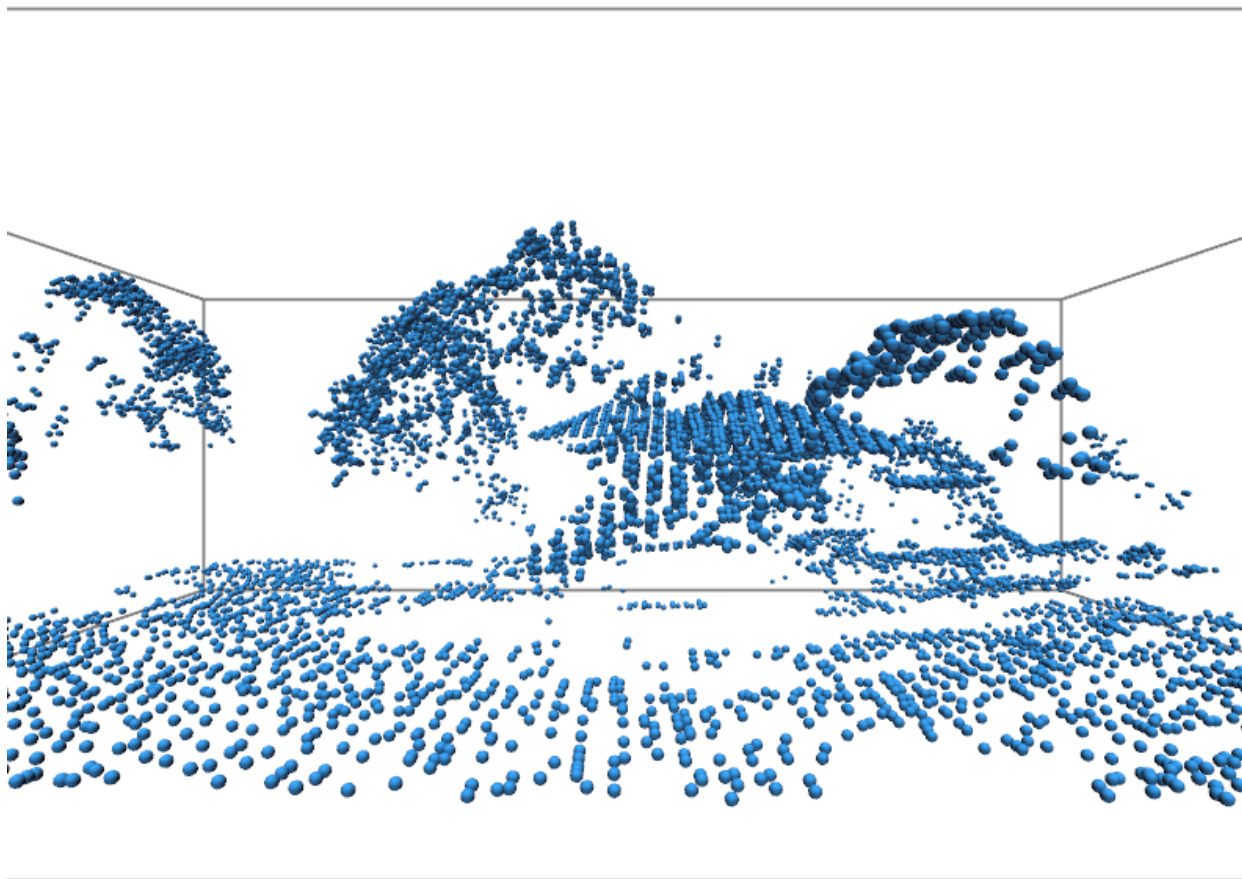
Plan view of Scene C, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (53m x 34m x 18m). Lidar data © Airbus Defence and Space Ltd. (2013b).

Scene C – North-facing

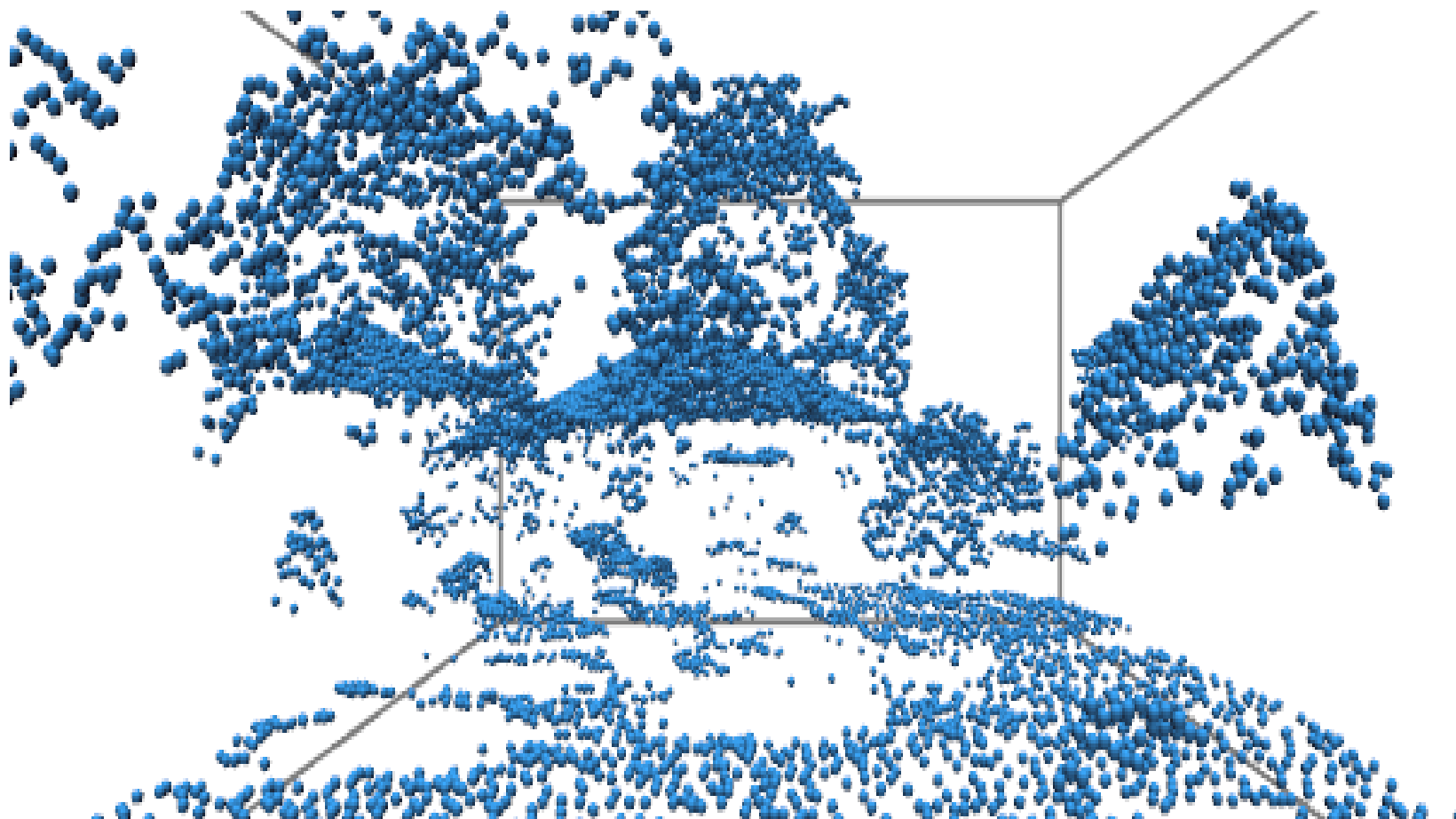
North-facing view of Scene C, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (53m x 34m x 18m). Lidar data © Airbus Defence and Space Ltd. (2013b).

Scene C – East-facing

East-facing view of Scene C, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (53m x 34m x 18m). Lidar data © Airbus Defence and Space Ltd. (2013b).

Scene C – South-facing

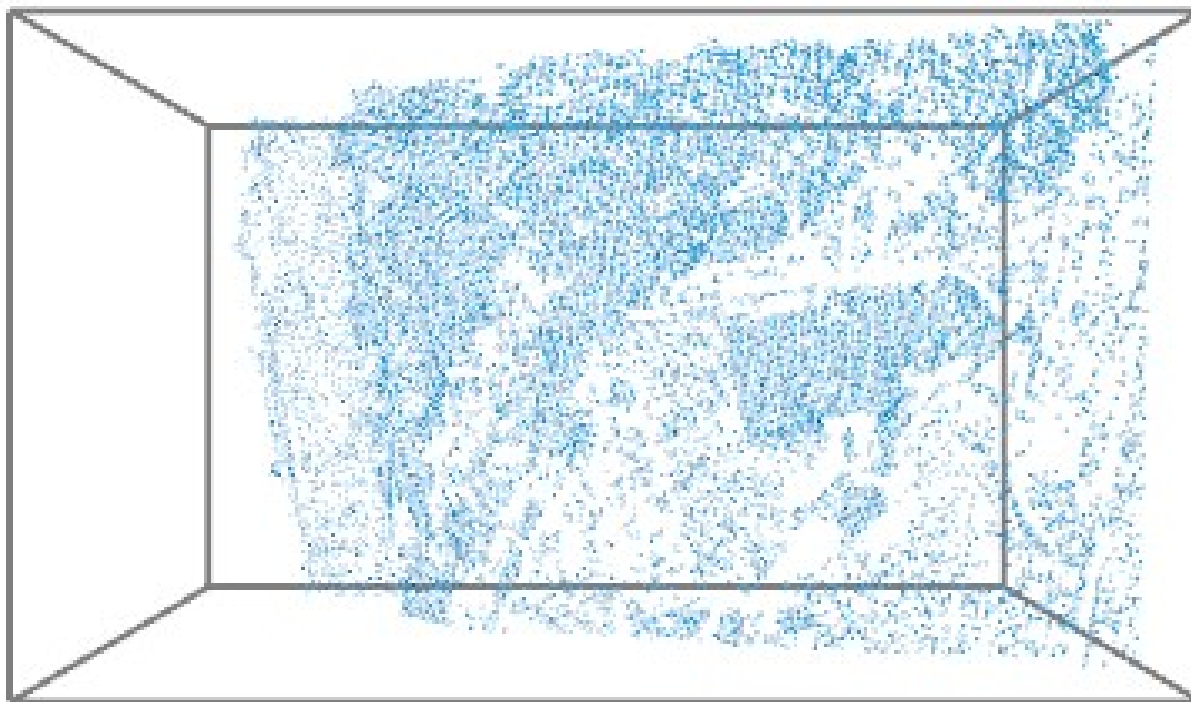
South-facing view of Scene C, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (53m x 34m x 18m). Lidar data © Airbus Defence and Space Ltd. (2013b).

Scene C – West-facing

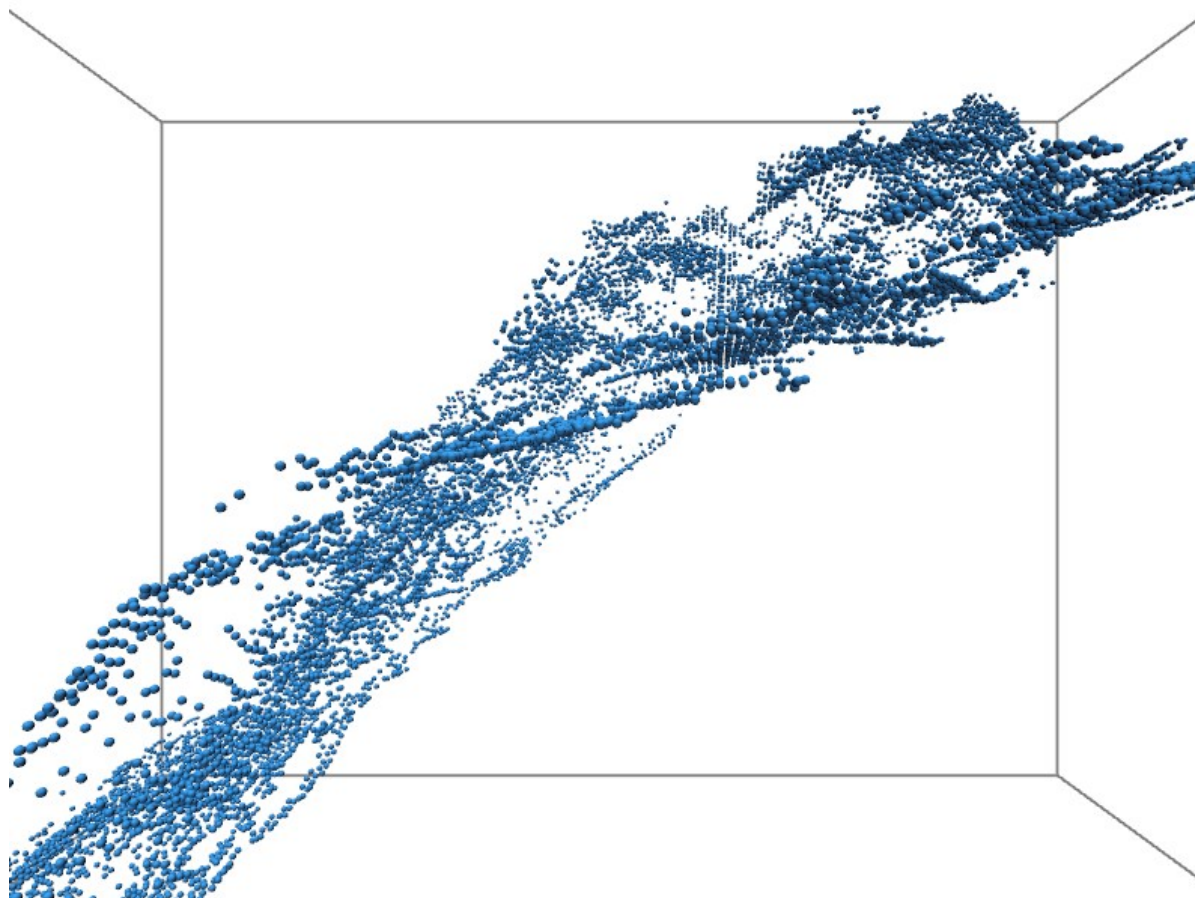
West-facing view of Scene C, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (53m x 34m x 18m). Lidar data © Airbus Defence and Space Ltd. (2013b).

Scene D – point cloud dataset

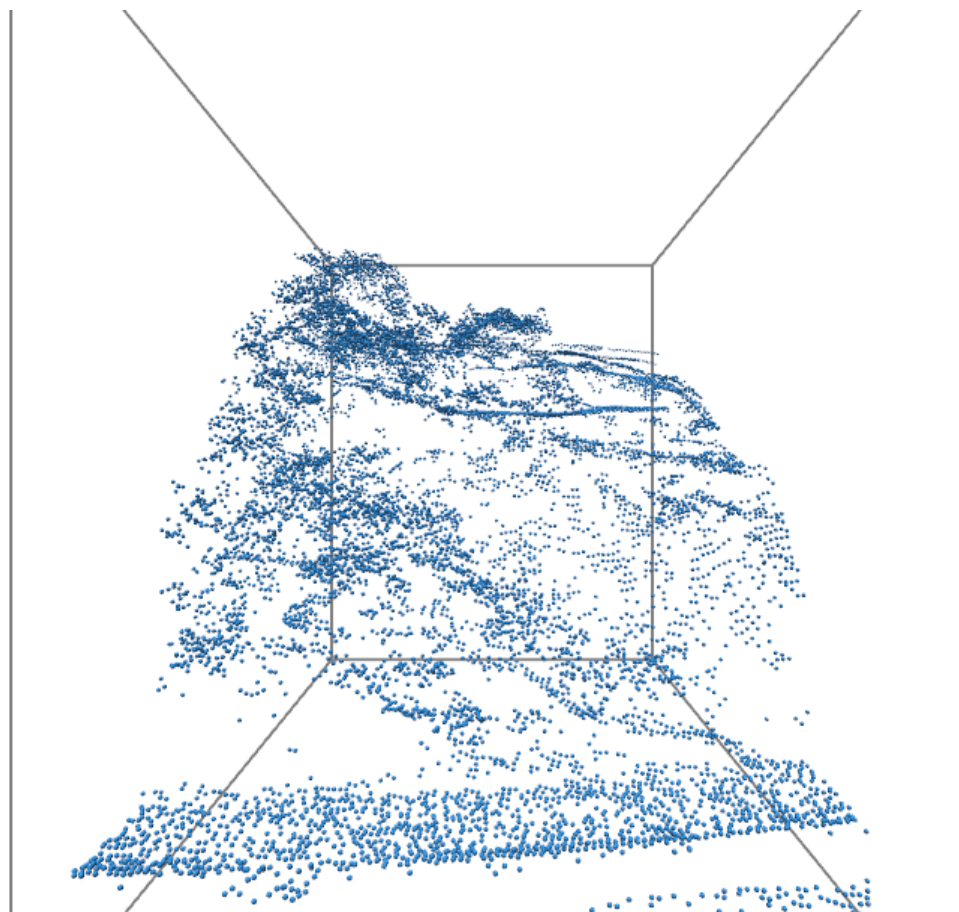
Plan view



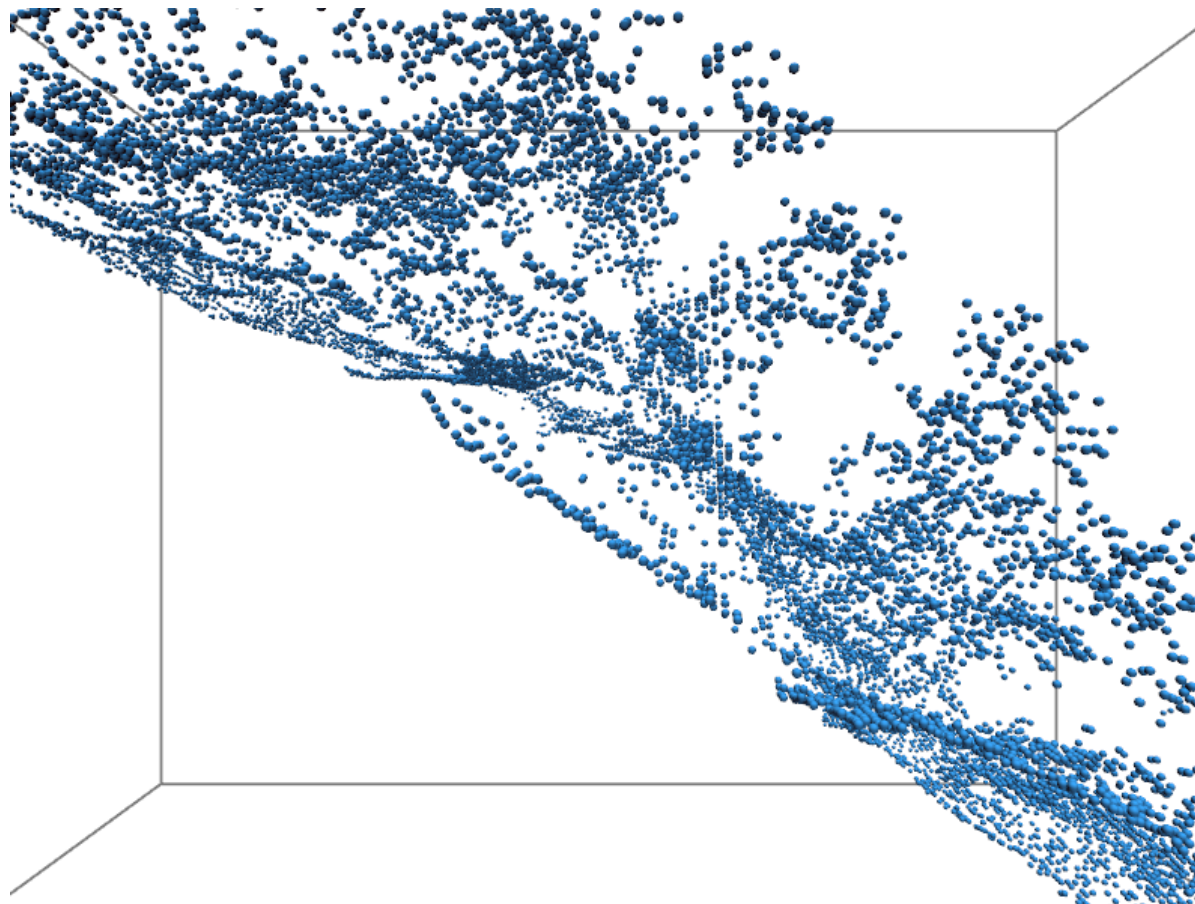
Plan view of Scene D, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (100m x 60m x 75m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene D – North-facing

North-facing view of Scene D, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (100m x 60m x 75m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene D – East-facing

East-facing view of Scene D, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (100m x 60m x 75m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Scene D – South-facing

South-facing view of Scene D, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (100m x 60m x 75m). Lidar data © Airbus Defence and Space Ltd. (2013a).

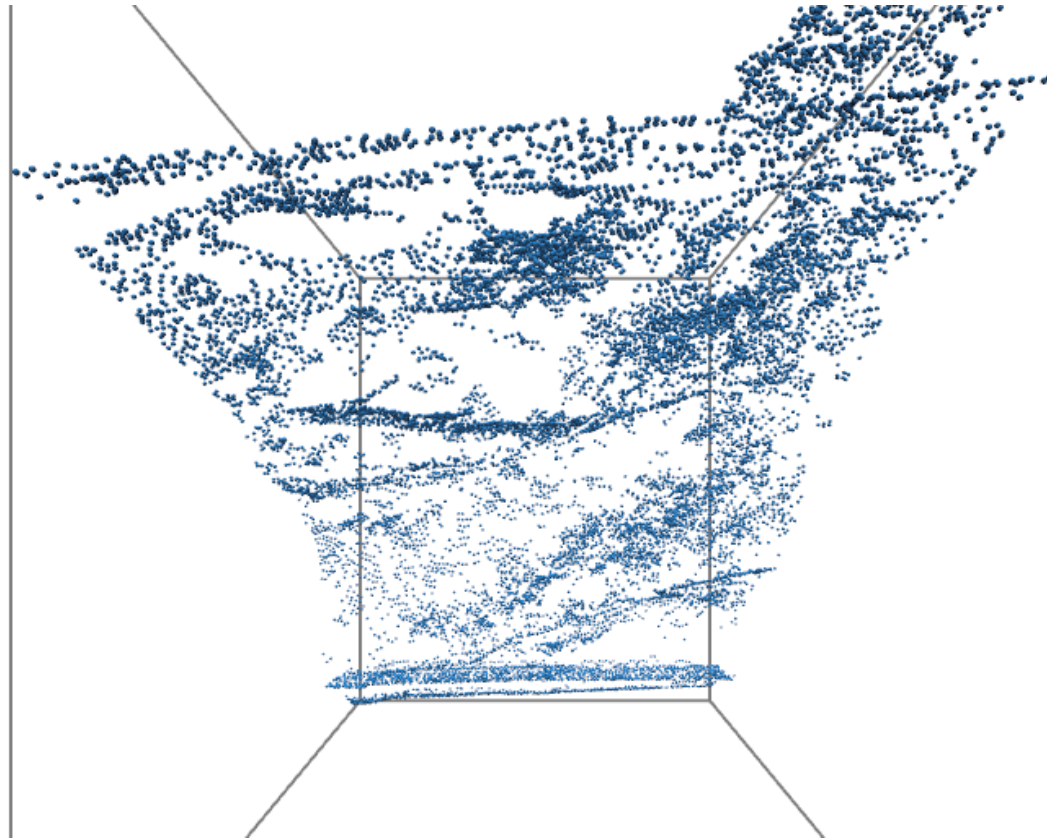
Scene D – West-facing

Figure 1. West-facing view of Scene D, which contains points of interest (POI) A to E (not highlighted, see Chapter 6 for detail). Scene was visualised in 2D or stereo 3D via Vizard 3.0 (WorldViz, 2010), displayed with a green background (RGB = 178, 223, 138) and white bounding box (100m x 60m x 75m). Lidar data © Airbus Defence and Space Ltd. (2013a).

Appendix D: Lessons learnt from pilot

| Participant | Amendments noted during session |
|-------------|---|
| PILOT1 | <ul style="list-style-type: none"> • Picture is needed to illustrate gamepad buttons to participant • Cover up alternative tests in Randot test to avoid distraction • Written survey – change ‘expert’ to ‘highly proficient’ • Gamepad configuration too sensitive • Participant recommends Scene C/D are not coloured by classification • Highlight points – turn off when looking at scene? • Change the colour of selected point during interpretation • For interpretation task, select point with gamepad, then describe POI • Colour mono and z value |
| PILOT2 | <ul style="list-style-type: none"> • Gamepad update [same configuration as above] • 8 cell grid during discussion of point cloud structure – long, but useful to PILOT2. • Colouring of POI is ambiguous – if participant does not name right colour (e.g. blue or purple?), don’t know which point they’re referring to • Could target point with trigger button and then describe POI verbally • Both PILOT1&2 state ground points first > put most POIs in the ambiguous areas • Turn off full screen view so window can minimize scene easily to look at participant instructions on the presentation slides • Set mouse colour to fade away when picked. |
| PILOT3 | <ul style="list-style-type: none"> • A/B scene overwrites output file. Do not write file each time – append to same file. • Tried out error selection with mouse. Must set mouse to disabled when picking with gamepad. • What is an “appropriate” ground point, when measuring? • Make sure colour of flashing points are distinctive against scene. • Participant had difficulty interpreting scene D (cliff) |
| PILOT4 | <ul style="list-style-type: none"> • Gamepad update – pilot gamepad code – had to nudge back to stop moving. <ul style="list-style-type: none"> - Tweak sensitivity of (1) move up+down, (2) chin up+down - Need a button to take user back to upright position. • Questionnaire. P4 stated never go to 3D cinema/sports, but commented that they had and it was like looking into a box as field of view was poor. • Change wording of questionnaire from ‘how regularly’ to ‘have you ever’. |

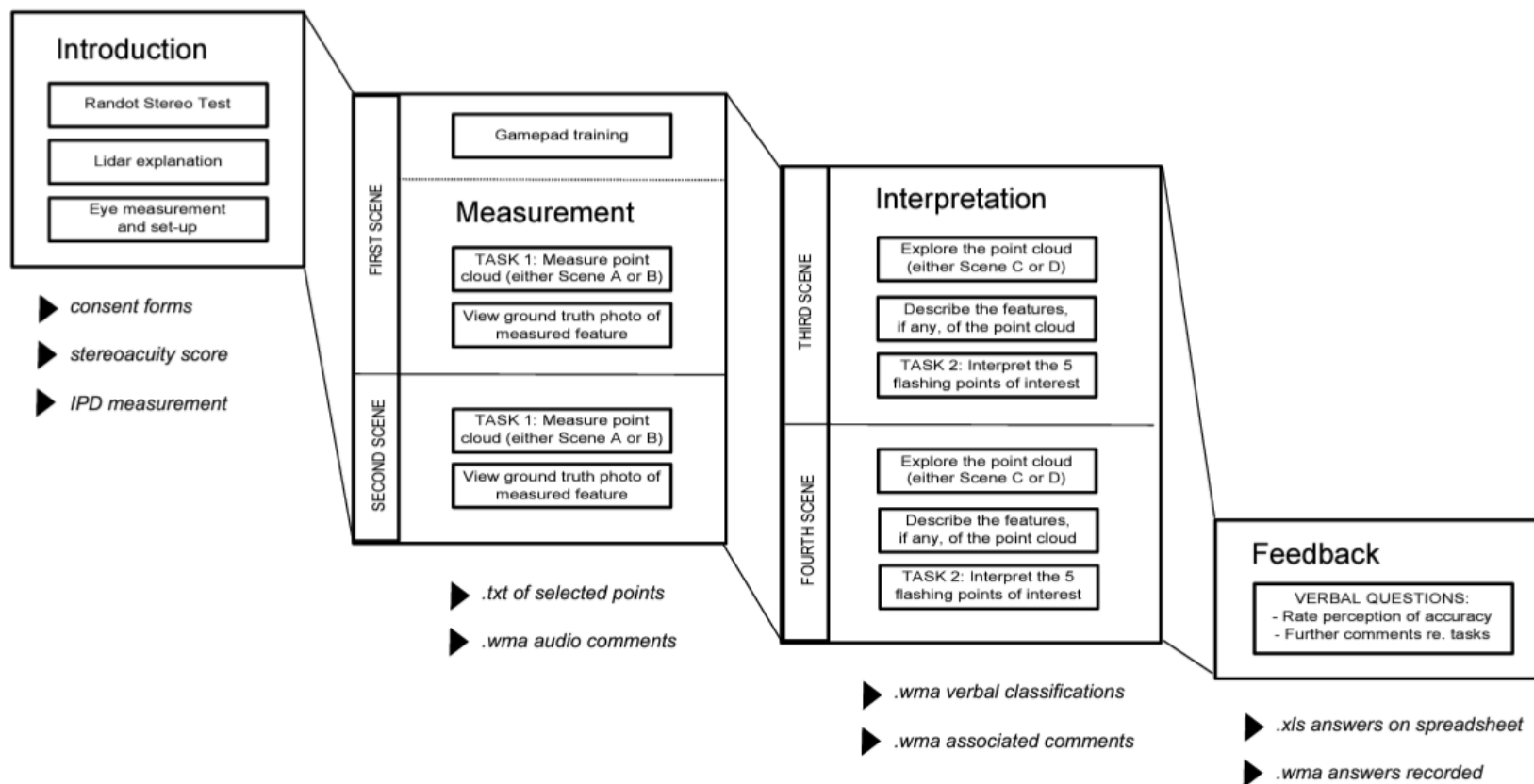
Appendix E: Display method per scene

Method used by each participant for Scenes A-D

| Pno | A | B | C | D |
|-----|----|----|----|----|
| P01 | NA | NA | 2D | 3D |
| P02 | NA | NA | NA | NA |
| P03 | NA | NA | 3D | 2D |
| P04 | 3D | 2D | 3D | 2D |
| P05 | NA | 3D | 2D | 3D |
| P06 | NA | NA | NA | NA |
| P07 | 2D | 3D | 2D | 3D |
| P08 | 3D | 2D | 3D | 2D |
| P09 | 2D | 3D | 2D | 3D |
| P10 | 3D | 2D | 3D | 2D |
| P11 | 2D | 3D | 2D | 3D |
| P12 | 3D | 2D | 3D | 2D |
| P13 | 3D | 2D | 3D | 2D |
| P14 | 2D | 3D | 2D | 3D |
| P15 | 3D | 2D | 3D | 2D |
| P16 | 2D | 3D | 2D | 3D |
| P17 | 2D | 3D | 2D | 3D |
| P18 | 2D | 3D | 2D | 3D |
| P19 | 3D | 2D | 3D | 2D |
| P20 | 3D | 2D | 3D | 2D |
| P21 | 2D | 3D | 2D | 3D |
| P22 | 3D | 2D | 3D | 2D |
| P23 | 3D | 2D | 3D | 2D |
| P24 | 2D | 3D | 2D | 3D |
| P25 | 2D | 3D | 2D | 3D |
| P26 | 2D | 3D | 2D | 3D |
| P27 | 2D | 3D | 2D | 3D |
| P28 | 3D | 2D | 3D | 2D |
| P29 | 3D | 2D | 3D | 2D |
| P30 | NA | NA | NA | NA |
| P31 | 3D | 2D | 3D | 2D |
| P32 | 2D | 3D | 3D | 2D |
| P33 | 3D | 2D | 3D | 2D |
| P34 | 3D | 2D | 3D | 2D |
| P35 | 3D | 2D | 3D | 2D |
| P36 | 3D | 2D | 3D | 2D |
| P37 | 2D | 3D | 2D | 3D |
| P38 | 3D | 2D | 3D | 2D |
| P39 | 2D | 3D | 2D | 3D |
| P40 | 2D | 3D | 2D | 3D |
| P41 | 2D | 3D | 2D | 3D |
| P42 | 2D | 3D | 2D | 3D |
| P43 | 3D | 2D | 3D | 2D |
| P44 | 2D | 3D | 2D | 3D |
| P45 | NA | NA | NA | NA |
| P46 | 2D | 3D | 2D | 3D |
| P47 | 3D | 2D | 3D | 2D |
| P48 | 3D | 2D | 3D | 2D |
| P49 | NA | NA | NA | NA |
| P50 | 2D | 3D | 2D | 3D |
| P51 | 3D | 2D | 3D | 2D |

| Scene Order | Randomised Participant Number | | | |
|-------------|-------------------------------|------|------|------|
| | 1 | 2 | 3 | 4 |
| 1st | A_2D | A_3D | B_2D | B_3D |
| 2nd | B_3D | B_2D | A_3D | A_2D |
| 3rd | C_2D | C_3D | D_2D | D_3D |
| 4th | D_3D | D_2D | C_3D | C_2D |

Appendix F: Flow diagram of participant experiment




Where ▶ denotes output results.

Appendix G: Participant documents

Information sheet

(page 1/2)

| | | | |
|---|------------------------------------|--|-------------|
|  | University of Leicester | <small>InfoSheet_BurwellGeogPhD_131114.doc</small> | Page 1 of 2 |
|---|------------------------------------|--|-------------|

Information Sheet

Title of the project: Using 2D and 3D vision to identify and measure geographical features inside a point cloud.

Principal Researcher: Claire Burwell, PhD student, Department of Geography.


Supervisors: Dr. Claire Jarvis & Dr. Kevin Tansey, Department of Geography.

Background to the study:

Purpose of the study: To find out how useful a 3D display of a geographical scene is compared to a 2D display of the same environment.

Method: Volunteers will be shown some small geographical areas that are represented by a cloud of points. These points are displayed via either: (a) a large flat 2D screen, or (b) a large screen that (when viewed using 3D glasses) displays the images in 3D.

An example of a 'point cloud' is shown below.



Why is my input needed for this project?
Volunteers are needed to help evaluate the 2D and 3D techniques.

What will the researcher do with my results?
Under each of the two techniques, volunteers' responses to different tasks are recorded. The results are used to help determine the pros and cons of applying 2D and 3D to this type of geographical data.

Your identity will not be disclosed.

If you would like to take this information away with you, please feel free to do so.

Information sheet (page 2/2)

InfoSheet_BurwellGeogPhD_131114.doc

Page 2 of 2

Inclusion criteria:

Can you see in stereo? A number of people are not able to see in 3D, so you will be tested to check whether you can or can't. Regrettably, if you can't see in 3D, you won't be able to take part in this particular experiment.

Are you sensitive to flickering images? Anyone who has a condition that can cause them to experience seizures or lose consciousness momentarily while viewing certain kinds of flashing lights or patterns cannot take part.

Other Information

1. **Eye measurements:** In addition to the stereo eye test, the distance between your eyes will be measured with a ruler. This is so that the 3D viewing settings can be tailored to each individual and to minimise visual discomfort.
2. **Recorded information:** Your on-screen actions (your location, what buttons you press) will be recorded via the computer and your verbal reactions will be recorded via an audio recording device (dictaphone).
3. **Location & Equipment:** The one-to-one session will take place in the Geography Department's virtual reality theatre, at:

*Room F55b, Geography Dept. Floor 1, Bennett Building,
University of Leicester, LE1 7RH.*

The room has a large screen, approx. 4m x 2.5m, and for half of the experiment you will be required to wear 3D glasses (these can be worn over normal glasses). Please note that the room will be dimly lit to allow for suitable viewing conditions.
4. **Experiment length:** The experiments last approximately 1 hour, but this will vary according to the individual. Although the test is timed, there are no time-limits for the test, so please give yourself extra time (in addition to the booked hour) in case we overrun.
5. **Visual Tasks:** You will be asked to take part in a series of tasks, which will involve you:
 - (a) *Looking at and exploring the displayed cloud of points.*
 - (b) *Identifying features from the scene (trees, buildings, etc.)*
 - (c) *Interacting with the point cloud and selecting points.*
6. **Timeframe:** Experiments will take place during Term 1. If you would like to take part, please email Claire Burwell, at clb50@le.ac.uk, to arrange a suitable date and time.
7. **Consent:** Once you have read this information sheet and the consent form, you will be asked to sign the consent form. By doing this, you confirm your understanding of your involvement in the study and agree to participate in the study.
8. **Access to data:** If you wish to find out about the results of this study, please send an email to Claire Burwell, at clb50@le.ac.uk.

If you would like to take this information away with you, please feel free to do so.

Consent form



ConsentForm_BurwellGeogPhD_131104.doc

Consent Form

Thank you for volunteering to take part in the following PhD study, which is being carried out by the Geography Department at the University of Leicester:

Title of the project: Using 2D and 3D vision to identify and measure geographical features inside a point cloud.

Principal Researcher: Claire Burwell, PhD student, Department of Geography.

PLEASE READ THROUGH THIS DOCUMENT CAREFULLY, MAKING SURE YOU AGREE TO ALL STATEMENTS BEFORE SIGNING THE BOTTOM OF THE PAGE.

Understanding of the Project

- I have had the project explained to me, and I have read the Information Sheet, which I may keep for my records.

I am happy...

- To have my 3D vision tested and my eyes measured.
- To be interviewed during the experiment.
- To allow an audio recording of the session.
- To have the results of my tests recorded by the computer so that 3D and 2D can be compared later by the researcher.
- To answer a short survey about my background, including answering questions about my age, gender, occupation, eyesight, and technological habits.



Data Protection:

This information will be held and used for the purpose of this research project.

- I understand that any information I provide is confidential and that no identifiable personal data will be published.
- I understand that no audio footage will be used for purposes beyond academic evaluation without my permission.
- I understand that I can request a transcription of my audio interview from the researcher.



Withdrawal

- I am aware and understand that I can withdraw from the study at any stage.

MY CONSENT: I am happy to take part in this project under the terms set out above.

Name Signature Date

Thank you for your consent to participate. If you have any questions or concerns about the nature of this assessment, please contact Claire Burwell at cib50@le.ac.uk.

Appendix H: Interviewer script

1. Introduction

Pre-recording

(1) Information Sheet

- Please read through if you haven't already
- If you have any questions, let me know.

(2) Consent form

- Once you are happy with the nature of the experiment, and want to continue, please read and complete the consent form.

Aim

- To understand how people interpret geographical data when using two different display types, 2-D and 3-D.
- The kind of data that we're looking at is remote-sensing data called lidar data.

Structure of experiment

1. Read over documents
2. Test of 3D vision
3. Background questions -verbal
4. Experiment
5. Feedback questions - verbal
6. Background questions - written


Vision


- Because this experiment does rely on your vision, I am going ask you a general question about your eyesight and do a 3-D vision test.


[Question below asked verbally]

Q) 2 • As far as you are aware, do you experience any problems with your vision? Say yes to all that apply.

- Short-sightedness (difficulties seeing things that are far from you)
- Far-sightedness (difficulties seeing things that are close to you)
- Colour blind (to any extent)
- Any other issues with my vision (please expand, if you wish)
- No problems that I am aware of
- I don't know

Q) 2.01  • If 'other issues', if you would like to explain further please do.

Q) 2.02  • If any problems stated,
"At the time of this experiment, are you wearing anything,
i.e. glasses or contact lenses, to correct your vision? "

 • If so which:

- Glasses
- Contact lenses
- None of the above

Stereo test

Reason 3D vision test required:

- During the experiment, you're going to be viewing this screen in 3D, via these glasses.
- They allow you to receive one image to one eye and a slightly different image to the other. In theory, this gives more depth to the display.
- This test will determine how well you can see in 3-D.
- If your results fall lower than a certain level, regrettably, you will not be able to take part in this particular study. I don't want you feel too disappointed if that is the case - about 1/10 people cannot see in stereo.

[Re-position lamp, to give ample light over the participant's shoulder]

PARTICIPANT INSTRUCTIONS:

- Wearing the glasses
- Hold book upright 16" from eye,
- Look at the circles.
- From 1 to 10, name which circle seems to float above the others - left, middle, or right.

"This gets increasingly harder as you go from 1 to 10."

| | | | | | | | | | |
|------------|------------|------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | | | | | | | | | |
| L | R | L | M | R | M | L | R | M | R |
| 400 | 200 | 140 | 100 | 70 | 50 | 40 | 30 | 25 | 20 |

OBSERVER INSTRUCTIONS:

- Record the level of stereopsis at the last one chosen correctly.
- If one is missed, go back and test the preceding line again to determine whether the subject can achieve this or is just guessing

End of stereo test

- >>6 [<<50 arcs]: I'm pleased to say your vision meets the requirements of the study so we will continue with the experiment.
- <<<5 [>50 arcs]: Unfortunately, your level of 3-D vision doesn't mean the required level, so you will not be able to take part in the rest of the study.

Lidar familiarity questions

Now that we are confident you can carry out the experiment, I would like to ask you some background questions about how familiar you are with the type of DATA we'll be viewing today.

Looking at the questions on the screen...

3.00) On a scale of 1 to 5, with 1 being not at all familiar and 5 being extremely familiar, how familiar are you with the following terms:

3.01) LiDAR

3.02) Laser-scanning

3.03)[If participant ranks 1 - this is a data collection method, which will be explained. If ranks 2 or higher:] "Could you please describe the term lidar/laser-scanning, to the best of your knowledge?"

4) How would you rate your level of lidar expertise, with one being none and five being highly proficient, compared to the average person in the street?

4.01) In months or years for how long would you say you have been familiar with lidar?

5.00) If you consider your level of expertise to be 2 or above, please describe your past experience with lidar data/laser-scanning

(research/project/module). What kind of features were in the scan? What scale? If you have processed the data was the point cloud, waveform, interpolated surface, DEM, values derived from lidar, acquired, process, viewed it, calculations, software?

Lidar overview

Everyone, regardless of their familiarity with lidar will be given the following overview of what we will be looking at today.

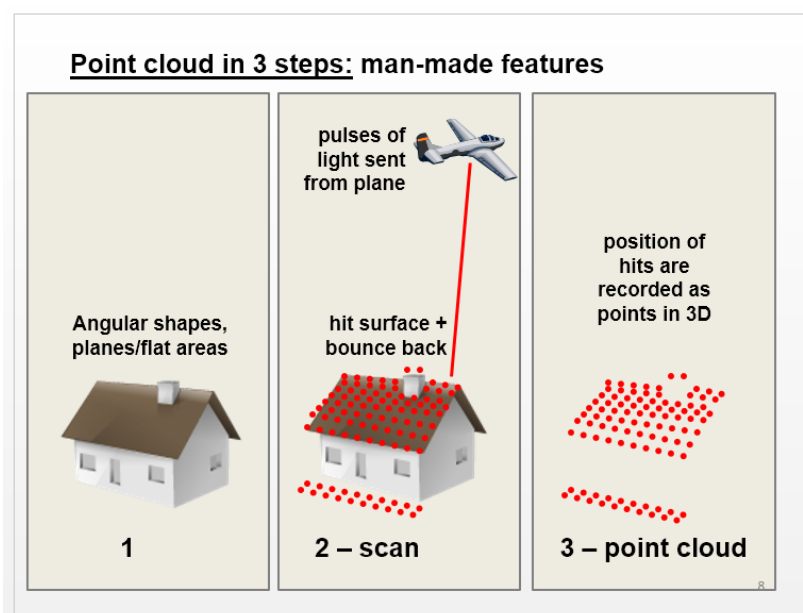


Image by Claire Burwell 2014

MAN-MADE FEATURE -

- 1) Generally more regular and angular in shape.
- 2) Scans of the environment taken from the air (e.g. aeroplane). Lasers are fired down towards the ground and bounce off surfaces that are hit.
- 3) The position of these hits make up a map of points that of the scanned environment, made up of points.
 - The lasers bounce off more solid features, like a roof or the ground. This might create an area of shadow where there are no points

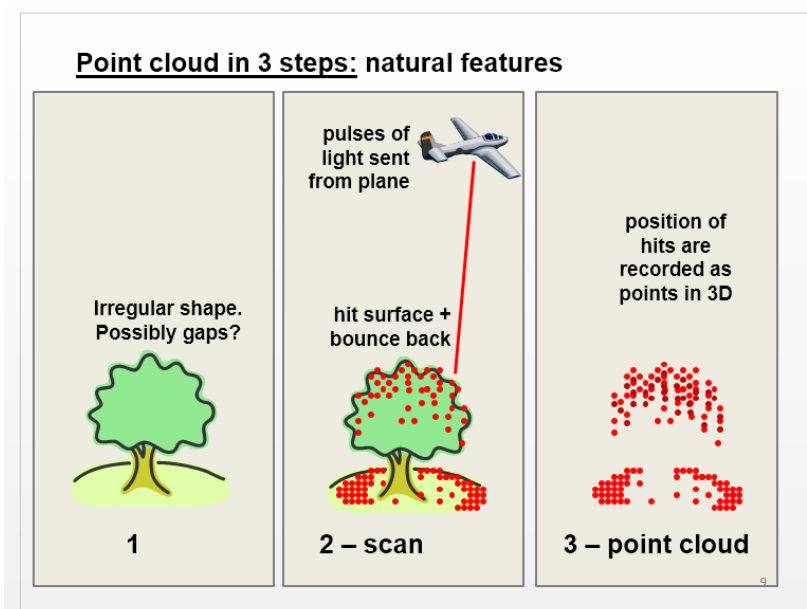


Image by Claire Burwell 2014

NATURAL FEATURES

- 1) Natural features, like trees, are more irregular.
- 2) Scans of the environment taken from the air (e.g. aeroplane). Lasers are fired down towards the ground and bounce off things that are hit.
- 3) The position of these hits make up a map of points that of the scanned environment, made up of points.
 - The light can penetrate through structures that have gaps, like some trees.

Do you have any questions about the overview?

Experiment procedure

- Real examples of point clouds are going to be presented to you.
- Four point clouds.
- Series of tasks, including recognition of features (trees, buildings, etc.) and measurement of the point cloud.
- I want to you always keep in mind that this is about you testing the methods, rather than me testing you.

Equipment Set-up

RE-SHUFFLE – You will be sat in this chair through the majority of the session; I'm going to be sat at this table towards the back of the room.

- Turn off the main lights [to reduce interference with the stereo projection], but this lamp at the back of the room will remain on.
- We will be looking at this large screen in front of us in 2D, like a normal

screen. When something is displayed in stereo, you will be able to see it in 3-D by wearing the 3-D glasses. I will hand a pair to you when you need them.

- Real examples of point clouds are going to be presented to you.
- NAUSEA I'm going to be showing the visualisations on the big screen and if you start to feel woozy or nauseous just let me know and we will stop.
- SOUND I'm going to move dictaphone a bit closer to you. You can hear that there is quite a noisy fan in the projector above, so I might sometimes ask you to just speak up either for the recorder or just because I'm a bit further away. A bit later on I might move in that chair when we're chatting more.

Measurement of interpupillary distance (IPD)

- Recall info sheet - need to measure the distance between your eyes.
- This just tailors the 3-D setting to you specifically, otherwise, it might look a bit funny or you might feel a bit queasy.
- I'm going to stand in front of you. If you keep both your eyes open, I'm going to put this ruler against your nose and then just read off your measurement.

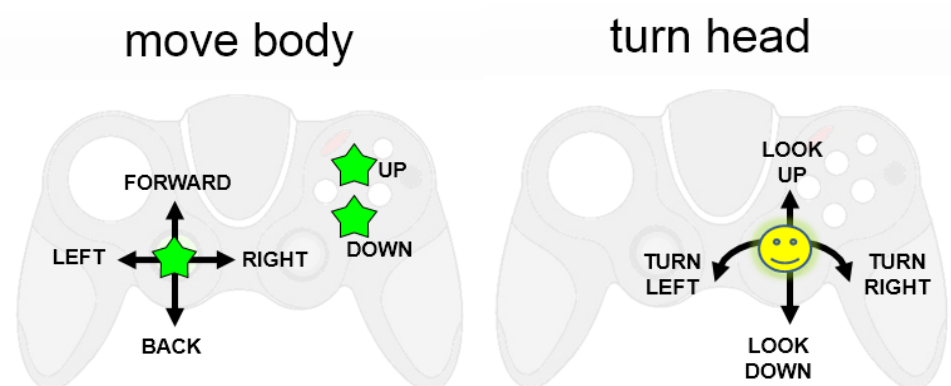
[inputs this value into the system, via GUI]

Sync warning

- GLASSES The large screen in front of us displays in 2D, like a normal screen. When something is displayed in stereo, you will be able to see it in 3D by wearing the 3D glasses.
- SYNC Sometimes the system doesn't synchronise properly with the glasses. The right eye should receive the right image, but sometimes the right image is show to the left eye and vice versa. When that happens, we have to wear the glasses upside down. I'll double check first whether we need to.
- We will begin by looking at Scene A/B

Gamepad Introduction

[The first scene is shown] You might normally interact with data on the computer with the mouse, but for this particular session I'm going to give you a gamepad to use. Do you ever use these? I'm just going to show you how it is configured because it might be different to what you are used to.



There are some stickers on the gamepad that correspond to the symbols on these images.

- To move your body/to move around the scene you use the left-hand joystick forward-backwards-left-right. To move up and down vertically, you used those two star buttons.
- To look around you use the right hand stick and to look up and down you go forwards.

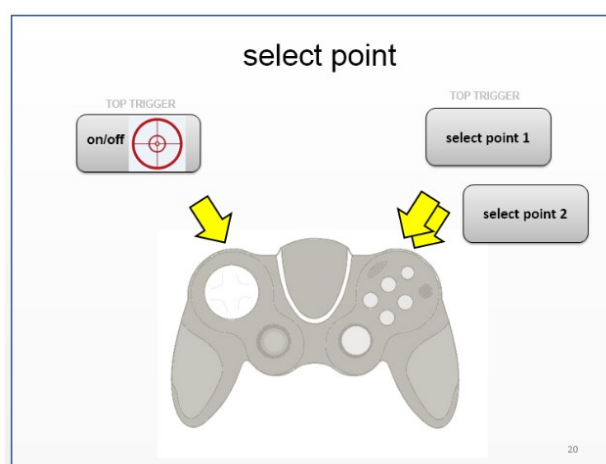
Explore & practise

- I'm going to take you back into the point cloud now, so please put your glasses back on [if necessary] and I will turn on the gamepad.
- Spend the next 30 seconds or so having a play with the gamepad and getting used to the controls.
- You are not being tested.

2. Measurement Task

Gamepad measuring

- This part of the experiment involves you measuring the features within the scenes, by selecting points.
- Shortly, I would like you to navigate around the scene and to take a measurement between two points. Effectively, you're selecting two points. In order to do that I'm going to introduce you to 3 more buttons. You will only use these buttons for this measurement task.



- LEFT The top left button will turn/toggle on and off the crosshair in the middle of the screen. That crosshair will remain in the centre of the screen throughout, so as you move is just going stay in the middle.
- RIGHT The buttons on the right hand side allow you to do the actual selection. When you have your first point within the crosshair, press the top button to select it. Use the bottom button when you get your second point, your endpoint.
- BUTTON NO. 3: If you want to flip yourself upright, you can press the number 3 button whenever you need to, to roll back upright.

TIPS DURING TASK:

- CROSS HAIR OFF: When navigating, make sure you have the target/cross-hair turned off, but when you want to select a point turn it on.
- JOYSTICK USE: Use the right-hand stick to move your head/ R stick to turn

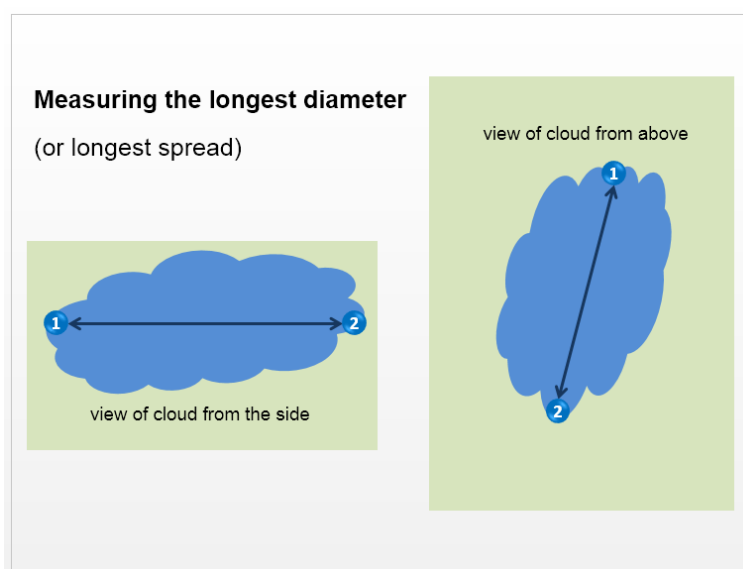
your head/turn head slower / touch the joystick very lightly.

- POSITION/ZOOM If you're having difficulty selecting the one you want, go up close up to select it, to make sure it's the one you want.

FIRST SCENE – A/B

- If 3D method used, check synchronisation of shutter glasses.
- It is up to you to have a look around and decide which two points you think are appropriate.

Measurement instructions for Scene B

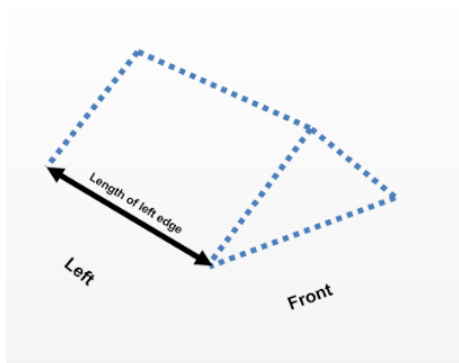


- QUOTE PARTICIPANT You mentioned <it looks like vegetation>.
- TASK Measure the longest diameter or spread of the feature.
e.g.
the width of the feature at its fullest part.
the two points where the feature is at its widest.
the two points that you consider to be the widest transect.

- VISUAL Show image on screen. Can go from any angle – images are examples only.

- (a) Select one point for your start
- (b) one for the end.

Measurement instructions for Scene A



- QUOTE PARTICIPANT You mentioned <it looks like a pitched roof>.
- TASK I'd like you to measure the lower left edge of the <feature/roof>
- VISUAL Show image on screen.

(a) Select one point for your start

(b) one for the end.

Context

- REAL WORLD I want to now show you this scene in real world context - I am going to show some images of what we've been looking at.
- If you would like to remove your glasses, I am going to reduce the screen.
- EXPECTED Is it different from what you expected? Is that the kind of thing you imagined?

SECOND SCENE - A/B

- If 3D method used, check synchronisation of shutter glasses.
- Repeat the task, using the appropriate measurement instructions

End of Part 1 / Break

- SUMMARY We've now looked at two small scenes - one man-made and one natural. They were an introduction to what an actual point cloud for those kind of different features.
- COMING UP In the next task, I'm going to show you 2 more scenes.
- Do you want to have a break? [Break according to participant preference]

3. Interpretation Task

- BIGGER AOI Now you will be looking at other, bigger, geographical scenes that are made up of a cloud of points.
- NO PHOTO You will not be shown any photographic images, so you'll be relying solely on the points.

THIRD SCENE – C or D

- If 3D method used, check synchronisation of shutter glasses.

Dimensions

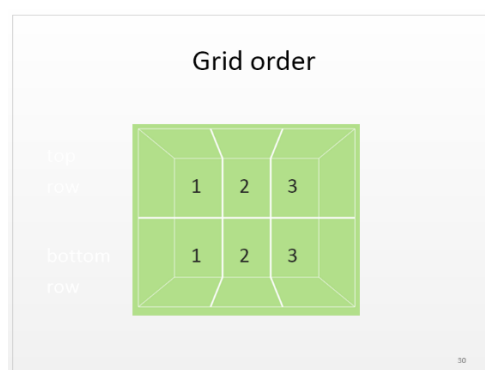
BOUNDING BOX This cube is a bounding box around the area that we are looking at. State dimensions. SHOW SCALE CUBE This represents 5 m³.

Initial look

Like the other two scenes, I am going to take us round each side. What are your initial thoughts?

Explore

- PRE-WARN - in a couple of minutes I will ask you to talk me through the scene. You will be given time to navigate around and familiarise yourself with the scene. Once you have had a look round, I will then take you back to a bird's eye view and will split the scene into six parts.



- SUMMARISE: If you want to just have a little look round for approx. 60 seconds, I will bring us back up to the top and I will ask you what features might be present in this scene.

Grid

- <prompt> Are you happy to go up to a bird's eye view in a moment?
- I am going to put a grid on the area.
- [add grid] Let me know what you think might be in each section, after your initial overview/Do you want to talk me through cell-by-cell?

POIs

- Now you have had an overview of the scene, I'd like you to see if you can identify what features the flashing points belong to.
- There are 5 points, across the scene.
- Search for the flashing points and decide: Does the point represent a tree, a house, etc.?
- You are encouraged to explore the scene to help you work this out.

FOURTH SCENE – C or D

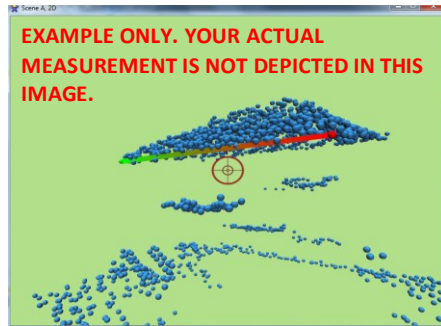
- If 3D method used, check synchronisation of shutter glasses.
- Repeat the steps above, for Third Scene.

4. Feedback Questions

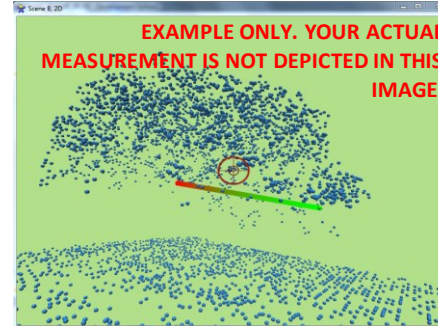
The following questions were shown on-screen, at the end of the experiment, and were discussed verbally. The interviewer recorded the answers directly into the Excel © (Microsoft Corporation, 2012) worksheet.

Q) 8.00 **MEASURING TASK**

e.g. Scene A measurement



e.g. Scene B measurement



0.01 How accurately do you feel you measured Scene A?

0.02 How accurately do you feel you measure Scene B?

Not at all accurately

Not at all accurately

1 2 3 4 5 6 7

| | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Very accurately

Very accurately

0.03 Any further comments?

TESTING 2D/3D WITH DIFFERENT SCENES

TYPE OF FEATURE

Q) 9 In general, how confident were you about identifying the following types of features, in the larger scenes?

| | | Very confident | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Not at all confident | n/a | Don't know |
|------|--|-------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|-------------------------|--------------------------|--------------------------|
| 0.01 | Natural features, in 2D | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | | <input type="checkbox"/> | <input type="checkbox"/> |
| 0.02 | Natural features, in 3D | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | | <input type="checkbox"/> | <input type="checkbox"/> |
| 0.03 | Man-made features, in 2D | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | | <input type="checkbox"/> | <input type="checkbox"/> |
| 0.04 | Man-made features, in 3D | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | | <input type="checkbox"/> | <input type="checkbox"/> |
| 0.05 | <input type="text"/> Any further comments? | | | | | | | | | | | |

TYPE OF SCENE

Q) 10 How confident were you about the answers you gave when identifying the feature of the flashing points ?

| | | Very confident | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Not at all confident |
|-------|------------------|-------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|-------------------------|
| 10.01 | Scene C - flat | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 10.02 | Scene D - sloped | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |

10.03 Any further comments?

Appendix I: Written survey



ID no. _____

WRITTEN SURVEY


TECHNOLOGY QUESTIONS

In this section we ask about your past experiences with different technologies

COMPUTER CONTROL DEVICES

Q) 13.00 How often do you play computer games with the following controls?

Consider frequency over the last year. If you used them previously, *i.e.* over 1 year ago, tick "In the past".

| | Never | Rarely (at least once per year) | Occasio- nally (at least once per month) | Often (at least once per week) | Very often (at least once per day) | In the past (1+ year ago) |
|--|--------------------------|--|--|---|--|---------------------------------|
| 13.01 Gamepad (for Playstation, Xbox, PC, etc. video games) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 13.02 Wiimote (for use with a Nintendo Wii games console) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 13.03 Kinect device (for use with an Xbox games console) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 13.04 Other navigation/control device(s) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 13.05  If any other(s), please name the device(s) here: | | | | | | |


13.06 Any further comments regarding this question:

DISPLAY TECHNOLOGIES

Q) 14.00 How often do you experience the following displays?

Consider frequency over the last year. If you used them previously, *i.e.* over 1 year ago, tick "In the past".

N.B. If your response is "Never" to ALL, go straight to BACKGROUND INFORMATION

| | Never | Rarely (at least once per year) | Occasio- nally (at least once per month) | Often (at least once per week) | Very often (at least once per day) | In the past (1+ year ago) |
|---|--------------------------|--|--|---|--|---------------------------------|
| 14.01 3DTV or 3D cinema screen (film, sporting event, etc.) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 14.02 3D stereoscopes (e.g. for paper-based stereo-mapping) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 14.03 3D stereo computer screen (e.g. digital stereo-mapping) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 14.04 Head-mounted display (e.g. Oculus Rift, etc.) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 14.05 Any other 3D displays (e.g. this VR Theatre, CAVE, etc.) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 14.06  If any other(s), please describe the display(s) here: | | | | | | |

14.07 Any further comments regarding this question:

Q) 15.00 If you have ever experienced any of the following 3D display technologies, how effective did you find them? 'Effective', meaning you felt that the 3D experience gave added depth to the images.

| | | Not at all effective (no extra depth in 3D) | ←————→ | | | Very effective (very strong 3D depth) | Never used (no comment, n/a) |
|-------|---|--|--------------------------|--------------------------|--------------------------|--|---------------------------------|
| 15.01 | 3DTV or 3D cinema screen (film, sporting event, etc.) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 15.02 | 3D stereoscopes (e.g. for paper-based stereo-mapping) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 15.03 | 3D stereo computer screen (e.g. digital stereo-mapping) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 15.04 | Head-mounted display (e.g. Oculus Rift, etc.) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 15.05 | Any other 3D displays (e.g. this VR Theatre, CAVE, etc.) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 15.06 | <i>If any other(s), please describe the display here:</i> | _____ | | | | | |
| 15.07 | Any further comments regarding this question: | _____ | | | | | |

BACKGROUND INFORMATION

These questions are posed to get an overview of the participants taking part in the study.

Q) 16 • Age

| | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 16-25 | 26-35 | 36-45 | 46-55 | 56+ |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Q) 17 • Gender

| | |
|--------------------------|--------------------------|
| Male | Female |
| <input type="checkbox"/> | <input type="checkbox"/> |

Q) 18 • What is your native language? *Please write:* _____

Q) 19 • What is your nationality? *Please write:* _____

Q) 20 • Occupation

| | | |
|--------------------------|--------------------------|--------------------------|
| Student | Staff | Other |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

↳ If you answered 'student', please circle your current...

Q) 20.01 • Degree type: BA BSc MA MSc PhD Other

If Other, please state.....

Q) 20.02 • Year of study: 1 2 3 3+

Q) 20.03 • Department (*please state*): _____

Q) 20.04 • Main subject area (*please state*): _____

↳ If you answered 'staff'

Q) 20.05 • In which department are you based?: _____

Q) 20.06 • What is your current role? _____

↳ If you answered 'other'

Q) 20.07 • Please state your current occupation: .. _____

Appendix J: Participant background

Lidar knowledge

List of participants' familiarity with the term 'lidar', 'laser-scanning', and ranking of their description of the term with which they are most familiar (if any). For description ranking, 0-3, where 0 = no knowledge of term(s), 3 = detailed and in-depth knowledge of described term.

| Participant number | Participant's ranking of their own familiarity with the term 'lidar' | Participant's ranking of their own familiarity with the term 'laser-scanning' | Assigned rank of participant description (of most familiar term) |
|--------------------|--|---|--|
| P01 | 5 = Extremely familiar | 4 = Moderately familiar | 3 |
| P02 | NA | NA | NA |
| P03 | 5 = Extremely familiar | 4 = Moderately familiar | 3 |
| P04 | 4 = Moderately familiar | 5 = Extremely familiar | 3 |
| P05 | 2 = Slightly familiar | 3 = Somewhat familiar | 2 |
| P06 | NA | NA | NA |
| P07 | 1 = Not at all familiar | 4 = Moderately familiar | 3 |
| P08 | 4 = Moderately familiar | 4 = Moderately familiar | 3 |
| P09 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P10 | 2 = Slightly familiar | 2 = Slightly familiar | 2 |
| P11 | 2 = Slightly familiar | 2 = Slightly familiar | 1 |
| P12 | 5 = Extremely familiar | 2 = Slightly familiar | 3 |
| P13 | 4 = Moderately familiar | 4 = Moderately familiar | 3 |
| P14 | 3 = Somewhat familiar | 3 = Somewhat familiar | 0 |
| P15 | 4 = Moderately familiar | 4 = Moderately familiar | 2 |
| P16 | 3 = Somewhat familiar | 2 = Slightly familiar | 0 |
| P17 | 5 = Extremely familiar | 5 = Extremely familiar | 1 |
| P18 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P19 | 4 = Moderately familiar | 4 = Moderately familiar | 3 |
| P20 | 4 = Moderately familiar | 4 = Moderately familiar | 2 |
| P21 | 4 = Moderately familiar | 4 = Moderately familiar | 3 |
| P22 | 4 = Moderately familiar | 4 = Moderately familiar | 3 |

| | | | |
|-----|-------------------------|-------------------------|----|
| P23 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P24 | 2 = Slightly familiar | 1 = Not at all familiar | 1 |
| P25 | 1 = Not at all familiar | 1 = Not at all familiar | 0 |
| P26 | 1 = Not at all familiar | 1 = Not at all familiar | 0 |
| P27 | 4 = Moderately familiar | 4 = Moderately familiar | 2 |
| P28 | 4 = Moderately familiar | 4 = Moderately familiar | 3 |
| P29 | 3 = Somewhat familiar | 2 = Slightly familiar | 3 |
| P30 | NA | NA | NA |
| P31 | 3 = Somewhat familiar | 2 = Slightly familiar | 2 |
| P32 | 2 = Slightly familiar | 2 = Slightly familiar | 1 |
| P33 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P34 | 2 = Slightly familiar | 2 = Slightly familiar | 2 |
| P35 | 4 = Moderately familiar | 4 = Moderately familiar | 2 |
| P36 | 3 = Somewhat familiar | 3 = Somewhat familiar | 0 |
| P37 | 4 = Moderately familiar | 2 = Slightly familiar | 0 |
| P38 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P39 | 3 = Somewhat familiar | 2 = Slightly familiar | 1 |
| P40 | 1 = Not at all familiar | 3 = Somewhat familiar | 1 |
| P41 | 1 = Not at all familiar | 2 = Slightly familiar | 2 |
| P42 | 3 = Somewhat familiar | 2 = Slightly familiar | 3 |
| P43 | 2 = Slightly familiar | 1 = Not at all familiar | 1 |
| P44 | 1 = Not at all familiar | 1 = Not at all familiar | 0 |
| P45 | NA | NA | NA |
| P46 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P47 | 4 = Moderately familiar | 4 = Moderately familiar | 2 |
| P48 | 5 = Extremely familiar | 5 = Extremely familiar | 3 |
| P49 | NA | NA | NA |
| P50 | 1 = Not at all familiar | 1 = Not at all familiar | 0 |
| P51 | 1 = Not at all familiar | 3 = Somewhat familiar | 1 |

Appendix K: Measurement results (RQ1)

Measurement estimates per participant – Scene A (2D & 3D)

The table below, *Meas_A.csv*, details the measurements made in Scene A and the results were used during analysis in R (R Core Team, 2014). NA denotes those readings that are not applicable, owing to stereovision results of $\leq 5/10$ or errors in file recordings. Regarding headers, *pno* = participant number, *randot* = Stereo test result (out of 10, where 10 = high stereoacuity), *secofarc* = stereo result in standard seconds of arc, *scene* = Scene A, *method* is 2D, 3D, or NA, *x1 y1 z1* refer to the xyz positions of first point selected, *x2 y2 z2* refer to second point, *xylen* is the distance between point 1 and 2 xy positions, *xyzlen* is the 3D distance between points 1 and 3, *scenefile* is the visualisation file (.py) used for that participant, *order* = scene shown first (1) or second (2).

| pno | randot | secofarc | scene | method | x1 | y1 | z1 | x2 | y2 | z2 | xylen | xyzlen | scenefile | order |
|-----|--------|----------|-------|--------|-------|-------|-------|-------|-------|-------|----------|----------|--------------|-------|
| P01 | 10 | 20 | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P02 | 0 | NA | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P03 | 7 | 40 | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P04 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P05 | 9 | 25 | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P06 | 0 | NA | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P07 | 9 | 25 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P08 | 7 | 40 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P09 | 10 | 20 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |
| P10 | 9 | 25 | A | 3D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P11 | 6 | 50 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |

| | | | | | | | | | | | | | | |
|-----|----|----|---|----|-------|-------|-------|-------|-------|-------|----------|----------|--------------|----|
| P12 | 10 | 20 | A | 3D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P13 | 6 | 50 | A | 3D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P14 | 8 | 30 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.76492 | 12.76869 | SceneA_2D_00 | 1 |
| P15 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P16 | 10 | 20 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |
| P17 | 9 | 25 | A | 2D | 12.13 | 12.29 | 11.08 | 12.13 | 12.29 | 11.08 | 12.94835 | 12.95308 | SceneA_2D_00 | 1 |
| P18 | 9 | 25 | A | 2D | 22 | 12.53 | 16.27 | 22 | 12.53 | 16.27 | 13.18214 | 13.18228 | SceneA_2D_00 | 2 |
| P19 | 8 | 30 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P20 | 7 | 40 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P21 | 9 | 25 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P22 | 8 | 30 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P23 | 8 | 30 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P24 | 7 | 40 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P25 | 9 | 25 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |
| P26 | 9 | 25 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 13.17367 | 13.17369 | SceneA_2D_00 | 2 |
| P27 | 10 | 20 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P28 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P29 | 10 | 20 | A | 3D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P30 | 5 | 70 | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P31 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P32 | 6 | 50 | A | 2D | 6.13 | 12.31 | 21.74 | 6.13 | 12.31 | 21.74 | 12.26296 | 12.26916 | SceneA_2D_00 | 2 |

| | | | | | | | | | | | | | | |
|-----|----|----|---|----|-------|-------|-------|-------|-------|-------|----------|----------|--------------|----|
| P33 | 7 | 40 | A | 3D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P34 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P35 | 10 | 20 | A | 3D | 6.13 | 12.31 | 21.74 | 6.13 | 12.31 | 21.74 | 12.23256 | 12.23258 | SceneA_3D_00 | 1 |
| P36 | 9 | 25 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P37 | 9 | 25 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.3789 | 12.38504 | SceneA_2D_00 | 1 |
| P38 | 7 | 40 | A | 3D | 6.32 | 11.75 | 20.78 | 6.32 | 11.75 | 20.78 | 11.3306 | 11.33188 | SceneA_3D_00 | 2 |
| P39 | 9 | 25 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P40 | 10 | 20 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P41 | 8 | 30 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |
| P42 | 7 | 40 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_2D_00 | 1 |
| P43 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 2 |
| P44 | 10 | 20 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.98 | 12.97886 | SceneA_2D_00 | 1 |
| P45 | 0 | NA | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P46 | 10 | 20 | A | 2D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |
| P47 | 9 | 25 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |
| P48 | 10 | 20 | A | 3D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 13.17367 | 13.17369 | SceneA_3D_00 | 2 |
| P49 | 0 | NA | A | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P50 | 9 | 25 | A | 2D | 5.79 | 11.94 | 22.37 | 5.79 | 11.94 | 22.37 | 12.97885 | 12.97886 | SceneA_2D_00 | 2 |
| P51 | 10 | 20 | A | 3D | 11.92 | 11.92 | 10.93 | 11.92 | 11.92 | 10.93 | 12.97885 | 12.97886 | SceneA_3D_00 | 1 |

Measurement estimates per participant – Scene B (2D & 3D)

The table below, *Meas_B.csv*, details the measurements made in Scene A and the results were used during analysis in R (R Core Team, 2014). NA denotes those readings that are not applicable, owing to stereovision results of $\leq 5/10$ or errors in file recordings. Regarding headers, *pno* = participant number, *randot* = Stereo test result (out of 10, where 10 = high stereoacuity), *secofarc* = stereo result in standard seconds of arc, *scene* = Scene A, method is 2D, 3D, or NA, *x1 y1 z1* refer to the xyz positions of first point selected, *x2 y2 z2* refer to second point, *xylen* is the distance between point 1 and 2 xy positions, *xyzlen* is the 3D distance between points 1 and 3, *scenefile* is the visualisation file (.py) used for that participant, *order* = scene shown first (1) or second (2).

| pno | randot | secofarc | scene | method | x1 | y1 | z1 | x2 | y2 | z2 | xylen | xyzlen | scenefile | order |
|-----|--------|----------|-------|--------|-------|-------|-------|-------|-------|-------|----------|----------|--------------|-------|
| P01 | 10 | 20 | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P02 | 0 | NA | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P03 | 7 | 40 | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P04 | 10 | 20 | B | 2D | 19.15 | 12.67 | 4.09 | 18.7 | 13.41 | 33.9 | 29.8134 | 29.82258 | SceneB_2D_00 | 1 |
| P05 | 9 | 25 | B | 3D | 5.45 | 10.53 | 11.64 | 38.27 | 14.66 | 23.95 | 35.05265 | 35.29512 | SceneB_3D_00 | 2 |
| P06 | 0 | NA | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P07 | 9 | 25 | B | 3D | 3.9 | 10.03 | 21.77 | 38.17 | 18.15 | 16.98 | 34.60313 | 35.54309 | SceneB_3D_00 | 2 |
| P08 | 7 | 40 | B | 2D | 5.45 | 10.53 | 11.64 | 37.21 | 13.62 | 10.87 | 31.76933 | 31.91925 | SceneB_2D_00 | 1 |
| P09 | 10 | 20 | B | 3D | 4.49 | 11.99 | 17.54 | 38.35 | 14.84 | 23.43 | 34.36847 | 34.48643 | SceneB_3D_00 | 1 |
| P10 | 9 | 25 | B | 2D | 38.17 | 18.15 | 16.98 | 4.08 | 9.66 | 20.7 | 34.29237 | 35.3277 | SceneB_2D_00 | 2 |
| P11 | 6 | 50 | B | 3D | 38.27 | 14.66 | 23.95 | 4.08 | 9.66 | 20.7 | 34.34412 | 34.70618 | SceneB_3D_00 | 1 |
| P12 | 10 | 20 | B | 2D | 3.9 | 10.03 | 21.77 | 35.06 | 22.01 | 16.2 | 31.65392 | 33.8451 | SceneB_2D_00 | 1 |
| P13 | 6 | 50 | B | 2D | 4.49 | 11.99 | 17.54 | 38.12 | 13.07 | 22.43 | 33.98366 | 34.00081 | SceneB_2D_00 | 2 |
| P14 | 8 | 30 | B | 3D | 4.08 | 9.66 | 20.7 | 38.47 | 13.15 | 13.55 | 35.12541 | 35.29837 | SceneB_3D_00 | 2 |

| | | | | | | | | | | | | | | |
|-----|----|----|---|----|-------|-------|-------|-------|-------|-------|----------|----------|--------------|----|
| P15 | 10 | 20 | B | 2D | 37.34 | 16.88 | 22.71 | 4.49 | 11.99 | 17.54 | 33.25434 | 33.61195 | SceneB_2D_00 | 2 |
| P16 | 10 | 20 | B | 3D | 3.9 | 10.03 | 21.77 | 38.35 | 14.84 | 23.43 | 34.48997 | 34.82376 | SceneB_3D_00 | 1 |
| P17 | 9 | 25 | B | 3D | 38.27 | 14.66 | 23.95 | 3.9 | 10.03 | 21.77 | 34.43907 | 34.7489 | SceneB_3D_00 | 2 |
| P18 | 9 | 25 | B | 3D | 3.9 | 10.03 | 21.77 | 38.35 | 14.84 | 23.43 | 34.48997 | 34.82376 | SceneB_3D_00 | 1 |
| P19 | 8 | 30 | B | 2D | 38.27 | 14.66 | 23.95 | 3.9 | 10.03 | 21.77 | 34.43907 | 34.7489 | SceneB_2D_00 | 1 |
| P20 | 7 | 40 | B | 2D | 6.66 | 16.71 | 22.91 | 38.17 | 18.15 | 16.98 | 32.06314 | 32.09546 | SceneB_2D_00 | 2 |
| P21 | 9 | 25 | B | 3D | 5.45 | 10.53 | 11.64 | 38.47 | 13.15 | 13.55 | 33.0752 | 33.1788 | SceneB_3D_00 | 2 |
| P22 | 8 | 30 | B | 2D | 3.9 | 10.03 | 21.77 | 38.35 | 14.84 | 23.43 | 34.48997 | 34.82376 | SceneB_2D_00 | 1 |
| P23 | 8 | 30 | B | 2D | 5.45 | 10.53 | 11.64 | 38.27 | 14.66 | 23.95 | 35.05265 | 35.29512 | SceneB_2D_00 | 1 |
| P24 | 7 | 40 | B | 3D | 38.27 | 14.66 | 23.95 | 3.9 | 10.03 | 21.77 | 34.43907 | 34.7489 | SceneB_3D_00 | 2 |
| P25 | 9 | 25 | B | 3D | 37.75 | 10.77 | 21.49 | 7.64 | 18.48 | 16.13 | 30.58336 | 31.54023 | SceneB_3D_00 | 1 |
| P26 | 9 | 25 | B | 3D | 5.45 | 10.53 | 11.64 | 38.35 | 14.84 | 23.43 | 34.94873 | 35.21349 | SceneB_3D_00 | 1 |
| P27 | 10 | 20 | B | 3D | 38.47 | 13.15 | 13.55 | 3.9 | 10.03 | 21.77 | 35.53383 | 35.67055 | SceneB_3D_00 | 2 |
| P28 | 10 | 20 | B | 2D | 38.35 | 14.84 | 23.43 | 5.45 | 10.53 | 11.64 | 34.94873 | 35.21349 | SceneB_2D_00 | 2 |
| P29 | 10 | 20 | B | 2D | 3.9 | 10.03 | 21.77 | 38.35 | 14.84 | 23.43 | 34.48997 | 34.82376 | SceneB_2D_00 | 2 |
| P30 | 5 | 70 | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P31 | 10 | 20 | B | 2D | 3.9 | 10.03 | 21.77 | 38.47 | 13.15 | 13.55 | 35.53383 | 35.67055 | SceneB_2D_00 | 1 |
| P32 | 6 | 50 | B | 3D | 5.45 | 10.53 | 11.64 | 38.17 | 18.15 | 16.98 | 33.15289 | 34.01732 | SceneB_3D_00 | 1 |
| P33 | 7 | 40 | B | 2D | 38.47 | 13.15 | 13.55 | 3.9 | 10.03 | 21.77 | 35.53383 | 35.67055 | SceneB_2D_00 | 1 |
| P34 | 10 | 20 | B | 2D | 38.47 | 13.15 | 13.55 | 7.54 | 12.06 | 28.69 | 34.43668 | 34.45392 | SceneB_2D_00 | 2 |
| P35 | 10 | 20 | B | 2D | 5.45 | 10.53 | 11.64 | 37.21 | 13.62 | 10.87 | 31.76933 | 31.91925 | SceneB_2D_00 | 2 |

| | | | | | | | | | | | | | | |
|-----|----|----|---|----|-------|-------|-------|-------|-------|-------|----------|----------|--------------|----|
| P36 | 9 | 25 | B | 2D | 4.08 | 9.66 | 20.7 | 38.35 | 14.84 | 23.43 | 34.37856 | 34.76662 | SceneB_2D_00 | 1 |
| P37 | 9 | 25 | B | 3D | 3.9 | 10.03 | 21.77 | 38.47 | 13.15 | 13.55 | 35.53383 | 35.67055 | SceneB_3D_00 | 2 |
| P38 | 7 | 40 | B | 2D | 24.59 | 22.26 | 13.52 | 24.11 | 11.08 | 22.36 | 8.853022 | 14.26073 | SceneB_2D_00 | 1 |
| P39 | 9 | 25 | B | 3D | 38.27 | 14.66 | 23.95 | 3.9 | 10.03 | 21.77 | 34.43907 | 34.7489 | SceneB_3D_00 | 2 |
| P40 | 10 | 20 | B | 3D | 5.45 | 10.53 | 11.64 | 38.27 | 14.66 | 23.95 | 35.05265 | 35.29512 | SceneB_3D_00 | 2 |
| P41 | 8 | 30 | B | 3D | 35.73 | 21.61 | 15.96 | 3.9 | 10.03 | 21.77 | 32.35591 | 34.3657 | SceneB_3D_00 | 1 |
| P42 | 7 | 40 | B | 3D | 3.9 | 10.03 | 21.77 | 38.27 | 14.66 | 23.95 | 34.43907 | 34.7489 | SceneB_3D_00 | 2 |
| P43 | 10 | 20 | B | 2D | 38.17 | 18.15 | 16.98 | 4.49 | 11.99 | 17.54 | 33.68465 | 34.24327 | SceneB_2D_00 | 1 |
| P44 | 10 | 20 | B | 3D | 5.45 | 10.53 | 11.64 | 38.27 | 14.66 | 23.95 | 35.05265 | 35.29512 | SceneB_3D_00 | 2 |
| P45 | 0 | NA | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P46 | 10 | 20 | B | 3D | 38.17 | 18.15 | 16.98 | 7.02 | 17.17 | 23.76 | 31.87932 | 31.89438 | SceneB_3D_00 | 1 |
| P47 | 9 | 25 | B | 2D | 35.2 | 11.8 | 7.78 | 7.54 | 12.06 | 28.69 | 34.67425 | 34.67523 | SceneB_2D_00 | 2 |
| P48 | 10 | 20 | B | 2D | 38.47 | 13.15 | 13.55 | 8.4 | 13.74 | 29.5 | 34.03832 | 34.04344 | SceneB_2D_00 | 1 |
| P49 | 0 | NA | B | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| P50 | 9 | 25 | B | 3D | 22.06 | 23.49 | 29.84 | 19.15 | 12.67 | 4.09 | 25.91391 | 28.08208 | SceneB_3D_00 | 1 |
| P51 | 10 | 20 | B | 2D | 38.47 | 13.15 | 13.55 | 5.45 | 10.53 | 11.64 | 33.0752 | 33.1788 | SceneB_2D_00 | 2 |

Measurement comments

N.B. Italicised text denotes interviewer.

Participant comments (2D and 3D) during anticipation of Scene A measurement task.

| Scene (A/B) | Method (2D/3D) | Order (1/2) | Participant | Quote (Response to, "What is your plan of action?", ahead of/during task) | Summary |
|-------------|----------------|-------------|-------------|---|--|
| A | 2D | 1 | P37 | 0:07:58.3 <i>You're saying you want to be more round, that way.</i> 0:08:00.4 Yeah, like parallel, you know what I mean. So I want to move round to that side, so that I can... | Perpendicular |
| A | 2D | 2 | P41 | 0:46:56.7 Well, positioned myself on the edge, where you want me to be, and then looking either top or down, so... there we go! So that's one. 0:47:31.3 That was easy< laughter> - DONE! 0:47:38.0 <i>How does it compare to the one before?</i> 0:47:34.7 Well, it's easy because it has the geometrical shape, so you just have to put yourself on the exact point of the shape that you want and in here even though there's no axis, because you know it's a oriented, you can put yourself parallel perpendicular to the axis of the edge and then move around that. | On the edge "easy" – geometrical shape provides axes |
| A | 2D | 2 | P25 | 0:51:56.0 Have a look from above, let's have a nose... Weeee! | From above |
| A | 2D | 2 | P26 | 0:54:02.3 I'm think I'm going to do it the same as before, from the side. So I'm looking at the left edge sort of straightforward. | Perpendicular |
| A | 3D | 2 | P04 | 34:08.8 I'm going to get all the roof in my view, square on, like the end I'm measuring. Then, I'm going to select my two furthest points and then go in and zap them. So, I chose my points and now I'm just gonna go and get them. | Perpendicular |
| A | 3D | 2 | P19 | P19 0:46:03.1 er...come here, try and get in the middle and then pick to see that... um.... [tails off, concentrating] | Perpendicular |

| | | | | | |
|---|----|---|-----|--|---------------------------|
| A | 3D | | | 1:04:28.4 I'm sort of trying to give myself a bit of distance this way, to see the full width... | Zoom out, full width |
| A | 3D | 1 | P47 | 0:39:25.4 I was gonna see what it looks like from the side first, along the length, and see if I could pick those points. Because I think I could see them quite clearly in that view. | Perpendicular, full width |
| A | 3D | 2 | P31 | 0:58:48.8 [After reiterating the instructions and restarting] I think I will. I just want to do a little bit from the top, instead of underneath. | From top |
| A | 3D | 2 | P43 | 1:00:03.3 [Plan of action is] the same as before, have a look from the top and this side. | Top, then side |
| A | 3D | 1 | P28 | 0:30:59.2 Um find the point which is furthest, closest to me in my current orientation and then when it's closest to this orientation. | Select closest |
| A | 3D | 1 | P29 | 0:36:43.1 How I'm planning to do it? Try to select the closest one to ME... and the two points furthest away on this left-hand... side... of the roof. | Select closest |

Participant comments (2D and 3D) during Feedback stage of the experiment, after rating difficulty of the measurement task for Scene A.

| Scene (A/B) | Method (2D/3D) | Order (1/2) | Participant | Quote (Reflective comments made during post-trial Feedback questions) | Summary |
|-------------|----------------|-------------|-------------|---|------------------------------|
| A | 2D | 2 | P11 | 0:40:23.3 [...] the house one.... they're much more clustered, so you couldn't see it quite as easily. | Clustered points |
| A | 2D | 1 | P21 | 1:29:51.0 [...] I still think I measured [B] accurately given that it was not, there wasn't as such clear boundaries as there were in Scene A? Because it was a roof of a building in Scene A obviously, so that's quite a distinct feature. | Clear boundaries Distinct |
| A | 2D | 2 | P25 | I think... it's easier to choose specific points on that one, because you're picking a flat edge, so you choose the edge. And you kind of choose the point that comes to the forefront of that edge... and you just go with whatever you sort of think on that. | Flat edge |

| | | | | | |
|---|----|---|-----|---|---|
| A | 3D | 2 | P31 | 0:46:27.9 Scene B, I didn't feel as accurate as Scene A, so scene B's like 5, while scene A I felt very accurate about it. Scene A was 3D, so I could easily detect the sides, the sides were very apparent. 0:47:23.6 Yes, I knew, in Scene A, I knew the sides, but where the edges came together [at the corners] was a little bit challenging, trying to depict. Because the point cloud even though the give the general sense of uniformity, it wasn't so... smooth as in to detect where [the] exact corner was. So, that was the challenge, but I could see this was the side. | Sides very apparent Corners had less uniformity. |
| A | 3D | 1 | P34 | 1:11:26.0 For Scene A I think I was very happy for my measurement for that. I wouldn't say 100%. It was helpful, the sort of STRAIGHT EDGE measurement. | Straight edge |
| A | 3D | 2 | P38 | 2:17:20.2 Scene A, it's easier for me to make out [than B in 2D]- this is a roof and this is a straight line on this roof. | Straight line |
| A | 3D | 2 | P43 | 0:41:05.7 but, um, Scene A I think was pretty much spot on because it was a nice straight line | Nice straight line |
| A | 3D | 2 | P48 | 0:49:26.8 For scene A, it's quite a linear feature, with a definite start and end point to it, so it was quite easy to sort of look at it and go, 'start measuring here' and 'start-'. It had a start and an end, essentially, and you could depict that quite easily. | Linear with start and end |

Participant comments (2D and 3D) during anticipation of Scene B measurement task.

| Scene (A/B) | Method (2D/3D) | Order (1/2) | Participant | Quote (Response to, "What is your plan of action?", ahead of/during task) | Summary |
|-------------|----------------|-------------|-------------|--|--|
| B | 2D | 2 | P07 | 0:37:31.2 <i>What are you trying to do, go over it?</i> 0:37:30.6 I was trying... well... I think if I go over it I will lose perspective on where the bottom one is, so is more trying to get just a better look at which part of it was widest. | Could lose perspective of ground points from above |

| | | | | | |
|---|----|----|-----|--|--|
| B | 2D | 2? | P13 | 0:39:57.0 I'm gonna have a look around first... to see if there's like... so see from both points of view? Like, around it, see if I can find the longest area. And then... | Look around first |
| B | 2D | 1 | P48 | 0:52:40.9 <i>You know how you looked around a bit and then you look to the top, what was the most useful?</i> 0:52:47.0 I think looking from the top-down because I felt when I was looking from the side, I was anything the points nearest to me and they were perhaps obscuring the points behind them that could have been wider, where is when you go to the top, I felt like naturally you would see the widest extent. | Plan – can see all points |
| B | 2D | 2 | P47 | 0:50:08.0 I'm thinking if I look at it in plan view, I'm not going to be able to distinguish the points... on the ground from the points in the cloud? Erm, or from these higher points. I'll have a look at it, but I don't think I'm going to be able to do it that way, unless they look larger... 0:50:43.3 Yeah, you can't, I don't think you can do it from plan view, think it's too... oh, maybe. 0:50:55.3 <i>What do you think now?</i> 0:50:54.6 If you zoom out, you can, because you sort of get, you get an impression of the shape of the points that are higher up, because they're... slightly larger? But, I'm not sure if I'd be CONFIDENT... that they were definitely not on the ground, if that makes sense 0:51:19.0 <i>So, what are your options then, what are you considering?</i> 0:51:27.6 I'm gonna look at it from, because I think... I get an impression of which might be the longest side from the plan view, and then I might use a cross-section view to actually pick the point out. | Plan in 2D – predict can't distinguish points. Zoom Not confident points are not on ground. Plan, then cross-section. |
| B | 2D | 1 | P31 | 0:44:24.2 it would be more challenging to do it from the top anyway, but let me look at it. Looking at it from the top, I think the challenge would be in, um.... The challenge from the top would be in confusing the base and the top of the cloud. 0:44:56.7 Looking at it from the side, may be quite easy because you can differentiate the base of the top of the feature, but if I turn the image on the top, I can - it's may be quite challenging. | Challenging from plan view, in 2D Confuse base and top |

| | | | | | |
|---|----|---|-----|---|--|
| B | 2D | 2 | P29 | <p>0:46:20.4 <i>What are you thinking?</i></p> <p>0:46:20.0 Yeah, from the top, it's difficult to see the difference. [...] The area where you have the highest density... this is might be where the object is to measure, for the difference between the... higher object and the ground. It's a little bit more difficult from the top... so I don't know if I can select the right point. Might be that... so will be a little bit bigger, only, so now... I think I have to do it from the outward... from the side view. [...] It helped to see where the... like the, it's a little bit oval-shaped, would you say oval? So then you can see which direction you have to measure, but which is the furthest point? I think better to see from, er, the side-view because I... I mix them up with the ground points [in plan view].</p> | View from top, mixed up ground and veg points, looked from side. |
| B | 2D | 2 | P28 | <p>0:48:54.6 First have a look from the top, I thought. Though, it's not that easy <laughter>.</p> | View from top, but hindered by navigation |
| B | 2D | 1 | P43 | <p>0:40:26.0 Um, I was going to have a look from above it can work out how to do that and then compare that to the side view, but then, you get lost in points.</p> | Plan to look from above, then compare to side |
| B | 2D | 2 | P51 | <p>1:06:00.6 First of all I want to, from downside, to see... but just, cause I want to ignore this floor things, but it's kind of... not, it's quite large when I look from downside, so I just prefer to look around may be how big this, and find the longest... and then make the point.</p> <p>1:06:28.1 <i>So you wanted to look from below, but?</i></p> <p>1:06:28.0 Yeah</p> <p>1:06:30.3 <i>because those points in the way</i></p> <p>1:06:35.8 Yes. It's a little bit hard if you just look from above. It's kind of... I can't recognise which is the floating one in this point, it's more difficult, but yeah... it's kind... as I find this longest, longest one and then I will go down straight. Right, so... and decide which point I want to choose.</p> | <p>Language – interviewer verifying description.</p> <p>Want to remove ground points that are obscuring view from below.</p> <p>From above, can't differentiate points</p> |
| B | 3D | 2 | P14 | <p>0:05:20.0 Going to see if I can look at it from the top.</p> | Top |

| | | | | | |
|---|----|---|-----|---|---|
| B | 3D | 2 | P44 | 0:43:19.4 I'll probably look from the top, at first, to see the longest diameter | Top first |
| B | 3D | 2 | P05 | 0:46:00.2 I'm gonna go- from this view I'm gonna go for this one [...] Obviously if I go around probably, from this view, probably from the other side would be different. No? Yeah. I'd say that's probably the longest axis. So from here, no, yeah I think that's the longest... yeah. | Predict views will be different |
| B | 3D | 2 | P42 | 1:01:13.9 <unprompted> I think being above is probably going to be helpful. 1:02:40.5 Yeah, I need to be able to see the 4 corners, don't I? A-ha! 1:02:52.3 <i>So, how are the 4 corners helping you?</i> 1:02:55.5 Well, it just occurred to me – if it's a scene and I'm looking down at it from above and you imagine that the scene's been taken looking down... everywhere it hits, I also need to be able to see, so I should in THEORY be able to see every point, except the ones that you know what going through the tree. Which... I should still be able to see, but, oh yeah! No, I should be able to see every point. 1:03:18.3 <i>Can you see them now, those points that you mentioned that have gone through the tree?</i> 1:03:21.8 Yeah, I can, and so if the tree was... you know the branches that are all kind of round here are obscuring my view of the corner... then obviously I'm not directly above, so obviously that's going to give me a perspective that's gonna throw me off trying to get the largest points. | From above [reason linked to acquisition] Can see ground points [no complaint of confusion with upper section] |
| B | 3D | 1 | P32 | 0:35:45.6 I'm planning to measure the distance, so I want to pick the best view so where I can- I expect that should be the easiest or the MOST accurate distance. 0:38:54.4 It's not accurate, but I think... 0:38:57.6 <i>Are you not happy?</i> 0:38:58.3 Yeah because there are some points at the back. If you look at the front view, you may not see those points, but when you zoom, then you can see the other points. So, based on the front view and from the side... I selected those points. | View from front occluded by points |

| | | | | | |
|---|----|---|-----|---|--------------------------|
| B | 3D | 2 | P27 | 0:40:19.1 I'm trying to get an overview of the entire... tree, so I can... estimate by eye which one, which is the longest axis and then choose the two points, which appear to be furthest away from each other. | overview |
| B | 3D | 2 | P21 | 0:39:01.3 Um, I'm looking at the image from the side to see which points seem to be furthest away in the scene? | Side |
| B | 3D | 2 | P07 | 0:38:51.7 <i>from this height, are you feeling the 3-D effect?</i> 0:38:57.4 Yeahhh, I still feel, well because I know the shape of it, I think it helps and also the dots that furthest away are smaller. | Plan |
| B | 3D | 1 | P41 | 0:35:12.4 Yes, I'm gonna go from the top, if it's the longest diameter, I've got to go from the top ... | Plan to go from the top. |

Participant comments (2D and 3D) during Feedback stage of the experiment, after rating difficulty of the measurement task for Scene B.

| Scene (A/B) | Method (2D/3D) | Order (1/2) | Participant | Quote (Reflective comments made during post-trial Feedback questions) | Summary |
|-------------|----------------|-------------|-------------|--|------------------------------|
| B | 3D | 1 | P25 | Looking from above is a bit easier, but, again, only if you're dead straight on to it. Um... and then you can't be 100% certain whether that point's...or, how high or how low it is in the canopy, if it's that. So, you can't judge height, so you're just really that, but then I guess that wasn't part of the... the thing. | |
| B | 2D | 1 | P04 | It [Scene B] was a bit harder to definitely make sure you got the edge branches. Just cause on the roof, where there's more points and more regular points, you could definitely make sure you got the corner. There was a corner to get, whereas on the tree, it was kind of judging which stuck out more - was a bit harder. | Judging which stuck out more |

| | | | | | |
|---|----|---|-----|---|--|
| B | 3D | 2 | P14 | <p>0:42:54.4 <i>so you did the first one in 2-D and the second one in 3-D. Because you said [lost from audio] you felt it was a bit less accurate for the 2nd scene. Was that more to do with this scene or more to do with the technique?</i></p> <p>0:43:06.7 I think it was just because of the scene, it was hard to tell what was the longest part.</p> | Scene rather than 2D/3D |
| B | 2D | 2 | P28 | <p>1:43:01.9 It was a bit more tricky because it wasn't such a structured thing, so it's hard to... I mean, even if one could physically half the thing, it might be also difficult to know exactly what's the... which one is the further diameter, yeah.</p> | Tricky |
| B | 2D | 2 | P29 | <p>1:33:45.4 The other one [B, in 2D] was more difficult [5] [...] It was not a regular shaped point cloud, so it...you have to look more precisely to these clouds.</p> | Shape means have to look more precisely |
| B | 2D | 1 | P47 | <p>0:49:55.2 Was that, that was the big misshapen cloud, wasn't it? Hmm [...] I think, um, it was an irregular shape, so actually having to work out where the 2 discrete end points were was a little bit harder to start with, compared to the roof shape.</p> <p>[b] P47 0:49:55.2 [...] Um... and... yeah, I... I think... the second one as well, you could, I think I spoke about this when I was doing it, but in PLAN view, you could work out ROUGHLY where the shape one, but then actually then determining between those points and the ground points, and being sure that they weren't ground points was a little bit trickier than the second one.</p> <p>0:51:19.2 <prompted> I think if it was in plan view, in 3D, and you'd got a bit more depth perception, then actually it might have been a little bit easier to do in 3D, on the second [scene B 2D] one. But I think it was just the irregular shape that really caused more problems! <laughter></p> | <p>Irregular shape</p> <p>Plan view</p> <p>Determining between points = tricky</p> <p>Irregular shape caused more problems (than 2D or 3D)</p> |
| B | 2D | 1 | P48 | <p>0:49:26.8 With the other scene [B], because it was a group of trees and it was essentially a mass of points, it didn't have anything, any uniformity to it or any linear features, so it was difficult to pick out a kind of natural starting endpoint when measuring it, so it was just a kind of case of looking at it in going I think</p> | <p>Difficult to pick out natural start and end.</p> <p>Estimating, not definite start and end.</p> |

| | | | | | |
|---|----|---|-----|--|---|
| | | | | these points of the 2 furthest apart almost, rather than going there's a definite start and a definite end. | |
| B | 3B | 2 | P24 | 1:23:50.5 [why gap in scores between A and B] The randomness of Scene B and the size...I think on that, it may have been easier with a mouse.[...] the sensitivity of the gamepad isn't as high as a sensitivity on the mouse. Although, using it for controlled navigation is easier, the sensitivity for selection is lower. | Randomness. Easier with mouse, gamepad sensitivity needs to be higher. |
| B | 3D | 1 | P38 | 2:17:20.2 This one, scene B, is more scattered than Scene A, so looking for the points, yeah. Scene A, it's easier for me to make out - this is a roof and this is a straight line on this roof. But, with scene B, it's scattered and I'm trying to look for the beginning and the end of one point. | Scattered Trying to look for the beginning and end |
| B | 3D | 2 | P40 | 1:22:11.7 5 it wasn't as straightforward. | Not straightforward |
| B | 3D | 2 | P27 | 1:21:41.2 it was difficult to work out which axis was the longest [in Scene B]. Once I'd worked out which axis was longest, it was comparatively easy to... get the longest points, but working it out was the hard part. | Working out = hard |

Further Measurement Accuracy Comments (Q8)

Further comments on Feedback Question 8, from participants who measured Scene B in 2D. Any comparisons made to 2D are in relation to Scene A.

| P. no. | 'Further comment' on Q8 - Measurement of Scene B (and participant used 2D method) | Task Order (1 st /2 nd) |
|--------|---|--|
| P04 | It was a bit harder to definitely make sure you got the edge branches | 1 st |
| P43 | 0:40:55.3 Um... I'd probably go for like 6. Pretty accurate but I might have got slightly the wrong point | 1 st |
| P34 | 1:11:54.8 For scene B though, I felt I wanted to change, obviously I wasn't able to, because I realised that it was slightly out. I measured it FAIRLY accurately, but not as accurately as Scene A. | 2 nd |
| P29 | 1:33:45.4 The other one [B, in 2D] was more difficult [5] [...] It was not a regular shaped point cloud, so it...you have to look more precisely to these clouds. | 2 nd |
| P36 | 1:36:14.7 [less confident in B (2D)] because when I navigate, the point doesn't seem as the farthest as I want it to. [...INT: Would it be fair to say then when you moved, it looked like it was wrong?] Yeah. Yes, this is what I felt, exactly. Which I even wanted to change the position of my point, the one I start with, but, yeah. I don't like it. | 1 st |
| P12 | 1:22:28.0 Scene B – I think I may have screwed that one up. Just when it [the line] went though, I felt like there was probably a longer diameter somewhere in there. [...] Just when I was trying to do it, I just couldn't really decide what stuff to do, because if I change the angle, then it would obviously look completely different. So, I like went around and decided that that point in the left on the bottom would be a good place to have the first one, but then deciding what the second one was was a bit haphazard. 1:23:18.6 [...] like I knew that that one would be a good one to off, but then it was just trying to go around and find the next one... 1:23:32.5 yeah, so that [first] one was clear, I guess because it's off, almost as a tail, and there wasn't a second one of that, so it was a bit. I wouldn't say I was massively off, but I wouldn't put any faith in mine being the longest diameter. | 1 st |

Further comments on Feedback Question 8, from participants who measured Scene B in 3D.

| P. no. | 'Further comment' on Q8 – Measurement of Scene B (and participant used 3D method) | Task Order (1 st /2 nd) |
|--------|--|--|
| P25 | 1:32:07.6 But, with Scene B... I think not very well at all. Because it was... erm... it had no flat edge , so you had to kind of try and create your own and... because there's no flat edge...'cause it's kind of like spherical, in a way. There's always gonna be another point that, when you look at it front-on, if you turn it a little bit MORE then there's ANOTHER outlier point and turn it again, there's ANOTHER out- so you constantly keep doing the same. Looking from above is a bit easier, but, again, only if you're dead straight on to it. Um... and then you can't be 100% certain whether that point's...or, how high or how low it is in the canopy, if it's that. So, you can't judge height. | 1 st |
| P26 | 1:36:44.4 I think I did quite well [Scene B]. | 1 st |
| P50 | 1:05:33.3 I couldn't navigate on scene B | 1 st |
| P14 | 0:43:06.7 I think it was just because of the scene , it was hard to tell what was the longest part. | 2 nd |
| P27 | 1:21:41.2 it was difficult to work out which axis was the longest in the 2 nd [scene]. Once I'd worked out which axis was longest, it was comparatively easy to... get the longest points, but working it out was the hard part. | 2 nd |
| P40 | 1:22:11.7 5 maybe because it wasn't as straightforward. | 2 nd |

Appendix L: Interpretation results (RQ2)

Summary of interpretation accuracy results

Where POI = point of interest, method values 2 and 3 = 2D and 3D, x = number of participants with correct feature interpretation for that POI, n = number of participants in sample, $p = x/n$.

| POI | scene | method | x | n | p |
|-----|-------|--------|----|----|-------|
| A | c | 2 | 16 | 22 | 0.727 |
| A | c | 3 | 18 | 24 | 0.750 |
| B | c | 2 | 14 | 22 | 0.636 |
| B | c | 3 | 22 | 24 | 0.917 |
| C | c | 2 | 8 | 22 | 0.364 |
| C | c | 3 | 9 | 24 | 0.375 |
| D | c | 2 | 21 | 22 | 0.955 |
| D | c | 3 | 24 | 24 | 1.000 |
| E | c | 2 | 9 | 22 | 0.409 |
| E | c | 3 | 16 | 24 | 0.667 |
| F | d | 3 | 13 | 22 | 0.591 |
| F | d | 2 | 17 | 24 | 0.708 |
| G | d | 2 | 17 | 24 | 0.708 |
| G | d | 3 | 21 | 22 | 0.955 |
| H | d | 2 | 18 | 24 | 0.750 |
| H | d | 3 | 21 | 22 | 0.955 |
| I | d | 2 | 11 | 24 | 0.458 |
| I | d | 3 | 11 | 22 | 0.500 |
| J | d | 2 | 10 | 24 | 0.417 |
| J | d | 3 | 10 | 22 | 0.455 |

2D & 3D interpretation for each POI in Scene C

Scene C; POI A: shrub

| partici- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | 2D | - | garden shed | - | 0 | - |
| P02 | NA | - | - | - | - | - | - |
| P03 | | - | 3D | - | Small tree | - | 1 |
| P04 | | - | 3D | - | small tree | - | 1 |
| P05 | | 2D | - | stairs | - | 0 | - |
| P06 | NA | - | - | - | - | - | - |
| P07 | | 2D | - | little tree | - | 1 | - |
| P08 | | - | 3D | - | small tree | - | 1 |
| P09 | | 2D | - | shrine (would say canopy, but b | - | 0 | - |
| P10 | | - | 3D | - | shed | - | 0 |
| P11 | | 2D | - | tree | - | 1 | - |
| P12 | | - | 3D | - | part of building | - | 0 |
| P13 | | - | 3D | - | small tree | - | 1 |
| P14 | | 2D | - | roof | - | 0 | - |
| P15 | | - | 3D | - | shrub | - | 1 |
| P16 | | 2D | - | small tree/branch/leaves | - | 1 | - |
| P17 | | 2D | - | small hedge/bush | - | 1 | - |
| P18 | | 2D | - | bush/tree | - | 1 | - |
| P19 | | - | 3D | - | unknown | - | 0 |
| P20 | | - | 3D | - | man-made | - | 0 |
| P21 | | 2D | - | small tree/bush | - | 1 | - |
| P22 | | - | 3D | - | bush/lower branch | - | 1 |
| P23 | | - | 3D | - | tree | - | 1 |
| P24 | | 2D | - | branch or leaf (vegetation) | - | 1 | - |
| P25 | | 2D | - | small tree/bush | - | 1 | - |
| P26 | | 2D | - | small tree or bush | - | 1 | - |
| P27 | | 2D | - | small tree | - | 1 | - |
| P28 | | - | 3D | - | small tree/shrub | - | 1 |
| P29 | | - | 3D | - | little bush or tree | - | 1 |
| P30 | NA | - | - | - | - | - | - |
| P31 | | - | 3D | - | tree, natural | - | 1 |
| P32 | | - | 3D | - | metal, man-made, [car], don't k | - | 0 |
| P33 | | - | 3D | - | small shrub | - | 1 |
| P34 | | - | 3D | - | low shrub | - | 1 |
| P35 | | - | 3D | - | small tree | - | 1 |
| P36 | | - | 3D | - | small tree | - | 1 |
| P37 | | 2D | - | ledge on the side of a building | - | 0 | - |
| P38 | | - | 3D | - | shrub | - | 1 |
| P39 | | 2D | - | tree | - | 1 | - |
| P40 | | 2D | - | shrubby | - | 1 | - |
| P41 | | 2D | - | small bush or small window or s | - | 1 | - |
| P42 | | 2D | - | bush | - | 1 | - |
| P43 | | - | 3D | - | vegetation | - | 1 |
| P44 | | 2D | - | mini-roof/statue | - | 0 | - |
| P45 | NA | - | - | - | - | - | - |
| P46 | | 2D | - | veg/bush | - | 1 | - |
| P47 | | - | 3D | - | shrub | - | 1 |
| P48 | | - | 3D | - | tree | - | 1 |
| P49 | NA | - | - | - | - | - | - |
| P50 | | 2D | - | vegetation | - | 1 | - |
| P51 | | - | 3D | - | small roof | - | 0 |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 22 | 24 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 16 | 18 |
| total incorrect | 6 | 6 |

Scene C; POI B: chimney

| particip- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|-------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | 2D | - | chimney or TV aerial | - | 0 | - |
| P02 | NA | - | - | - | - | - | - |
| P03 | | - | 3D | - | Chimney or [man-made] structure | - | 1 |
| P04 | | - | 3D | - | chimney | - | 1 |
| P05 | | 2D | - | chimney | - | 1 | - |
| P06 | NA | - | - | - | - | - | - |
| P07 | | 2D | - | velux window | - | 0 | - |
| P08 | | - | 3D | - | chimney | - | 1 |
| P09 | | 2D | - | chimney | - | 1 | - |
| P10 | | - | 3D | - | chimney | - | 1 |
| P11 | | 2D | - | chimney or aerial | - | 1 | - |
| P12 | | - | 3D | - | chimney [with bit of vegetation | - | 1 |
| P13 | | - | 3D | - | chimney | - | 1 |
| P14 | | 2D | - | chimney | - | 1 | - |
| P15 | | - | 3D | - | chimney | - | 1 |
| P16 | | 2D | - | chimney | - | 1 | - |
| P17 | | 2D | - | don't know, not connected to ro | - | 0 | - |
| P18 | | 2D | - | big branch | - | 0 | - |
| P19 | | - | 3D | - | chimney | - | 1 |
| P20 | | - | 3D | - | chimney | - | 1 |
| P21 | | 2D | - | sky window in roof | - | 0 | - |
| P22 | | - | 3D | - | chimney | - | 1 |
| P23 | | - | 3D | - | chimney | - | 1 |
| P24 | | 2D | - | chimney | - | 1 | - |
| P25 | | 2D | - | [dormer] window | - | 0 | - |
| P26 | | 2D | - | chimney | - | 1 | - |
| P27 | | 2D | - | chimney or something else stick | - | 1 | - |
| P28 | | - | 3D | - | chimney | - | 1 |
| P29 | | - | 3D | - | chimney, though irregularly sha | - | 1 |
| P30 | NA | - | - | - | - | - | - |
| P31 | | - | 3D | - | [after doubt introduced] chimne | - | 1 |
| P32 | | - | 3D | - | chimney | - | 1 |
| P33 | | - | 3D | - | chimney | - | 1 |
| P34 | | - | 3D | - | chimney pot or aerial | - | 1 |
| P35 | | - | 3D | - | roof | - | 0 |
| P36 | | - | 3D | - | chimney | - | 1 |
| P37 | | 2D | - | chimney | - | 1 | - |
| P38 | | - | 3D | - | chimney | - | 1 |
| P39 | | 2D | - | chimney | - | 1 | - |
| P40 | | 2D | - | chimney | - | 1 | - |
| P41 | | 2D | - | chimney | - | 1 | - |
| P42 | | 2D | - | aerial/bird | - | 0 | - |
| P43 | | - | 3D | - | chimney | - | 1 |
| P44 | | 2D | - | chimney | - | 1 | - |
| P45 | NA | - | - | - | - | - | - |
| P46 | | 2D | - | chimney | - | 1 | - |
| P47 | | - | 3D | - | chimney | - | 0 |
| P48 | | - | 3D | - | chimney | - | 1 |
| P49 | NA | - | - | - | - | - | - |
| P50 | | 2D | - | chimney/plant | - | 0 | - |
| P51 | | - | 3D | - | chimney | - | 1 |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 22 | 24 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 14 | 22 |
| total incorrect | 8 | 2 |

Scene C; POI C: hedge

| particip- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|-------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | 2D | - | hedge | - | 1 | - |
| P02 | NA | - | - | - | - | - | - |
| P03 | | - | 3D | - | wall with plants on | - | 1 |
| P04 | | - | 3D | - | hedge or wall | - | 0 |
| P05 | | 2D | - | fence covered by shrubs, someth | - | 1 | - |
| P06 | NA | - | - | - | - | - | - |
| P07 | | 2D | - | hedge | - | 1 | - |
| P08 | | - | 3D | - | man-made construction - buildin | - | 0 |
| P09 | | 2D | - | artificial - | - | 0 | - |
| P10 | | - | 3D | - | hedge | - | 1 |
| P11 | | 2D | - | path - grass or path | - | 0 | - |
| P12 | | - | 3D | - | hedgerow or stonewall | - | 1 |
| P13 | | - | 3D | - | wall | - | 0 |
| P14 | | 2D | - | stairway/ramp | - | 0 | - |
| P15 | | - | 3D | - | pipes | - | 0 |
| P16 | | 2D | - | Tunnel > wall/fence | - | 0 | - |
| P17 | | 2D | - | wall/large fence, although not | - | 0 | - |
| P18 | | 2D | - | no answer (but "dense") | - | 0 | - |
| P19 | | - | 3D | - | fence | - | 0 |
| P20 | | - | 3D | - | hedge | - | 1 |
| P21 | | 2D | - | covered walkway | - | 0 | - |
| P22 | | - | 3D | - | small row of bushes | - | 1 |
| P23 | | - | 3D | - | formal hedge or partly overgrow | - | 1 |
| P24 | | 2D | - | vehicle or roof of a walkway | - | 0 | - |
| P25 | | 2D | - | trellis/covered walkway | - | 0 | - |
| P26 | | 2D | - | veg on wall | - | 1 | - |
| P27 | | 2D | - | hedge | - | 1 | - |
| P28 | | - | 3D | - | little wall | - | 0 |
| P29 | | - | 3D | - | wall | - | 0 |
| P30 | NA | - | - | - | - | - | - |
| P31 | | - | 3D | - | hedges/ flowers that have been | - | 1 |
| P32 | | - | 3D | - | path with plants above path | - | 0 |
| P33 | | - | 3D | - | wall | - | 0 |
| P34 | | - | 3D | - | man-made raised walkway | - | 0 |
| P35 | | - | 3D | - | corridor | - | 0 |
| P36 | | - | 3D | - | small wall | - | 0 |
| P37 | | 2D | - | raised natural feature or man-m | - | 0 | - |
| P38 | | - | 3D | - | line of trees | - | 1 |
| P39 | | 2D | - | end of hedge | - | 1 | - |
| P40 | | 2D | - | hedge | - | 1 | - |
| P41 | | 2D | - | ground | - | 0 | - |
| P42 | | 2D | - | rhododendron bush | - | 1 | - |
| P43 | | - | 3D | - | hedge/fence/wall | - | 0 |
| P44 | | 2D | - | wall | - | 0 | - |
| P45 | NA | - | - | - | - | - | - |
| P46 | | 2D | - | buidling | - | 0 | - |
| P47 | | - | 3D | - | hedgerow | - | 1 |
| P48 | | - | 3D | - | covered walkway | - | 0 |
| P49 | NA | - | - | - | - | - | - |
| P50 | | 2D | - | man-made cover to walkway | - | 0 | - |
| P51 | | - | 3D | - | man-made walkway / path with co | - | 0 |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 22 | 24 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 8 | 9 |
| total incorrect | 14 | 15 |

Scene C; POI D: extension

| particip- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|-------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | 2D | - | lowered bit of roof | - | 1 | - |
| P02 | NA | - | - | - | - | - | - |
| P03 | | - | 3D | - | Extension Roof | - | 1 |
| P04 | | - | 3D | - | extension flat roof | - | 1 |
| P05 | | 2D | - | flat roof | - | 1 | - |
| P06 | NA | - | - | - | - | - | - |
| P07 | | 2D | - | out-house | - | 0 | - |
| P08 | | - | 3D | - | extension roof | - | 1 |
| P09 | | 2D | - | lower flat roofline or porch | - | 1 | - |
| P10 | | - | 3D | - | porch roof | - | 1 |
| P11 | | 2D | - | part of the house, man-made | - | 1 | - |
| P12 | | - | 3D | - | extension of building | - | 1 |
| P13 | | - | 3D | - | extension roof | - | 1 |
| P14 | | 2D | - | roof/building | - | 1 | - |
| P15 | | - | 3D | - | extension | - | 1 |
| P16 | | 2D | - | top part of building where it g | - | 1 | - |
| P17 | | 2D | - | building roof | - | 1 | - |
| P18 | | 2D | - | flat, lower roof | - | 1 | - |
| P19 | | - | 3D | - | roof of a porch | - | 1 |
| P20 | | - | 3D | - | addition to building | - | 1 |
| P21 | | 2D | - | part of building | - | 1 | - |
| P22 | | - | 3D | - | extension | - | 1 |
| P23 | | - | 3D | - | projecting bit of the house | - | 1 |
| P24 | | 2D | - | flat roof | - | 1 | - |
| P25 | | 2D | - | flat roof | - | 1 | - |
| P26 | | 2D | - | roof of extension / over window | - | 1 | - |
| P27 | | 2D | - | roof of lower section | - | 1 | - |
| P28 | | - | 3D | - | extension | - | 1 |
| P29 | | - | 3D | - | part of house (roof/balconey/fl | - | 1 |
| P30 | NA | - | - | - | - | - | - |
| P31 | | - | 3D | - | garage/porch, extending, not fr | - | 1 |
| P32 | | - | 3D | - | roof | - | 1 |
| P33 | | - | 3D | - | flat roof to extension | - | 1 |
| P34 | | - | 3D | - | flat-roofed extension | - | 1 |
| P35 | | - | 3D | - | roof | - | 1 |
| P36 | | - | 3D | - | roof | - | 1 |
| P37 | | 2D | - | lower ledge of building | - | 1 | - |
| P38 | | - | 3D | - | something attached to the house | - | 1 |
| P39 | | 2D | - | flat roof, off house | - | 1 | - |
| P40 | | 2D | - | flat roof | - | 1 | - |
| P41 | | 2D | - | man-made structure (entrance/wi | - | 1 | - |
| P42 | | 2D | - | add-on to building | - | 1 | - |
| P43 | | - | 3D | - | flat roof | - | 1 |
| P44 | | 2D | - | part of roof | - | 1 | - |
| P45 | NA | - | - | - | - | - | - |
| P46 | | 2D | - | roof of bits that stick out | - | 1 | - |
| P47 | | - | 3D | - | another building adjacent to ho | - | 1 |
| P48 | | - | 3D | - | extension/building | - | 1 |
| P49 | NA | - | - | - | - | - | - |
| P50 | | 2D | - | roof/extension | - | 1 | - |
| P51 | | - | 3D | - | roof | - | 1 |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 22 | 24 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 21 | 24 |
| total incorrect | 1 | 0 |

Scene C; POI E: tree

| particip- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|-------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D inepretation | 3D interpretation | 2D | 3D |
| P01 | | 2D | - | roof or tree | - | 0 | - |
| P02 | NA | - | - | - | - | - | - |
| P03 | | - | 3D | - | roof | - | 0 |
| P04 | | - | 3D | - | branch | - | 1 |
| P05 | | 2D | - | edge of roof | - | 0 | - |
| P06 | NA | - | - | - | - | - | - |
| P07 | | 2D | - | corner of roof | - | 0 | - |
| P08 | | - | 3D | - | roof | - | 0 |
| P09 | | 2D | - | roof or vegetation | - | 0 | - |
| P10 | | - | 3D | - | corner of roof | - | 0 |
| P11 | | 2D | - | roof (just!) | - | 0 | - |
| P12 | | - | 3D | - | tree | - | 1 |
| P13 | | - | 3D | - | tree | - | 1 |
| P14 | | 2D | - | roof corner | - | 0 | - |
| P15 | | - | 3D | - | tree | - | 1 |
| P16 | | 2D | - | tree | - | 1 | - |
| P17 | | 2D | - | tree | - | 1 | - |
| P18 | | 2D | - | edge of roof OR tree | - | 0 | - |
| P19 | | - | 3D | - | tree | - | 1 |
| P20 | | - | 3D | - | tree | - | 1 |
| P21 | | 2D | - | part of a tree | - | 1 | - |
| P22 | | - | 3D | - | roof/tree | - | 0 |
| P23 | | - | 3D | - | tree | - | 1 |
| P24 | | 2D | - | corner of roof | - | 0 | - |
| P25 | | 2D | - | natural [tree?] | - | 1 | - |
| P26 | | 2D | - | tree | - | 1 | - |
| P27 | | 2D | - | tree | - | 1 | - |
| P28 | | - | 3D | - | tree | - | 1 |
| P29 | | - | 3D | - | tree | - | 1 |
| P30 | NA | - | - | - | - | - | - |
| P31 | | - | 3D | - | tree by side of house | - | 1 |
| P32 | | - | 3D | - | edge of roof | - | 0 |
| P33 | | - | 3D | - | corner of building | - | 0 |
| P34 | | - | 3D | - | part of tree abutting the house | - | 1 |
| P35 | | - | 3D | - | roof | - | 0 |
| P36 | | - | 3D | - | tree | - | 1 |
| P37 | | 2D | - | roof | - | 0 | - |
| P38 | | - | 3D | - | branch of a tree | - | 1 |
| P39 | | 2D | - | corner of roof of house | - | 0 | - |
| P40 | | 2D | - | roof | - | 0 | - |
| P41 | | 2D | - | roof | - | 0 | - |
| P42 | | 2D | - | one of 2 trees | - | 1 | - |
| P43 | | - | 3D | - | tree | - | 1 |
| P44 | | 2D | - | tree | - | 1 | - |
| P45 | NA | - | - | - | - | - | - |
| P46 | | 2D | - | bottom of roof intermingled wit | - | 0 | - |
| P47 | | - | 3D | - | tree | - | 1 |
| P48 | | - | 3D | - | tree | - | 1 |
| P49 | NA | - | - | - | - | - | - |
| P50 | | 2D | - | part of tree | - | 1 | - |
| P51 | | - | 3D | - | corner of roof | - | 0 |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 22 | 24 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 9 | 16 |
| total incorrect | 13 | 8 |

2D & 3D interpretation for each POI in Scene D

Scene D; POI F: road

| partici- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | - | 3D | - | boardwalk or elevated road | - | 1 |
| P02 | NA | - | - | - | - | - | - |
| P03 | | 2D | - | Road | - | 1 | - |
| P04 | | 2D | - | man-made flat ledge/ level (NOT | - | 1 | - |
| P05 | | - | 3D | - | at the edge of a paved road | - | 1 |
| P06 | NA | - | - | - | - | - | - |
| P07 | | - | 3D | - | road | - | 1 |
| P08 | | 2D | - | road/path | - | 1 | - |
| P09 | | - | 3D | - | road/path | - | 1 |
| P10 | | 2D | - | terrace (a rice-field terrace o | - | 0 | - |
| P11 | | - | 3D | - | flat, like a terrace, NOT man-m | - | 0 |
| P12 | | 2D | - | part of pathway/road | - | 1 | - |
| P13 | | 2D | - | ground | - | 0 | - |
| P14 | | - | 3D | - | terrace/platform | - | 0 |
| P15 | | 2D | - | grassy non-mountainous area | - | 0 | - |
| P16 | | - | 3D | - | road | - | 1 |
| P17 | | - | 3D | - | flat part of ground | - | 0 |
| P18 | | - | 3D | - | road/debris flow | - | 1 |
| P19 | | 2D | - | river terrace | - | 1 | - |
| P20 | | 2D | - | man-made flat ground | - | 1 | - |
| P21 | | - | 3D | - | road | - | 1 |
| P22 | | 2D | - | paved ground | - | 1 | - |
| P23 | | 2D | - | artificial terrace | - | 1 | - |
| P24 | | - | 3D | - | man-made floor | - | 1 |
| P25 | | - | 3D | - | sea-wall or road at base of cli | - | 1 |
| P26 | | - | 3D | - | road | - | 1 |
| P27 | | - | 3D | - | Man-made, flattened. Terraces (| - | 0 |
| P28 | | 2D | - | road | - | 1 | - |
| P29 | | 2D | - | street | - | 1 | - |
| P30 | NA | - | - | - | - | - | - |
| P31 | | 2D | - | road or small footbal field or | - | 1 | - |
| P32 | | 2D | - | ground surface ("not ground-gro | - | 0 | - |
| P33 | | 2D | - | something like a railway suppor | - | 1 | - |
| P34 | | 2D | - | road or terrace | - | 1 | - |
| P35 | | 2D | - | vegetation | - | 0 | - |
| P36 | | 2D | - | vegetation or land | - | 0 | - |
| P37 | | - | 3D | - | ground, start or mountain | - | 0 |
| P38 | | 2D | - | ground | - | 0 | - |
| P39 | | - | 3D | - | flat area before a drop of clif | - | 0 |
| P40 | | - | 3D | - | like a shelf, ground overhangin | - | 0 |
| P41 | | - | 3D | - | road | - | 1 |
| P42 | | - | 3D | - | ledge, bank, flat surface at gr | - | 0 |
| P43 | | 2D | - | road | - | 1 | - |
| P44 | | - | 3D | - | bottom of hill, sea-wall | - | 0 |
| P45 | NA | - | - | - | - | - | - |
| P46 | | - | 3D | - | path/road | - | 1 |
| P47 | | 2D | - | road/track surface | - | 1 | - |
| P48 | | 2D | - | road | - | 1 | - |
| P49 | NA | - | - | - | - | - | - |
| P50 | | - | 3D | - | road | - | 1 |
| P51 | | 2D | - | road | - | 1 | - |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 24 | 22 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 17 | 13 |
| total incorrect | 7 | 9 |

Scene D; POI G: cliff

| particip- ant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|------------------|--------|-----|-----|---|----------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | - | 3D | - | bit of cliff | - | 1 |
| P02 | NA | - | - | - | - | - | - |
| P03 | | 2D | - | waterfall | - | 0 | - |
| P04 | | 2D | - | grassy bank | - | 1 | - |
| P05 | | - | 3D | - | ground and grass | - | 1 |
| P06 | NA | - | - | - | - | - | - |
| P07 | | - | 3D | - | slope | - | 1 |
| P08 | | 2D | - | ground | - | 1 | - |
| P09 | | - | 3D | - | cliff face/slope | - | 1 |
| P10 | | 2D | - | slope | - | 1 | - |
| P11 | | - | 3D | - | hill, natural undulating surface | - | 1 |
| P12 | | 2D | - | man-made support/arch | - | 0 | - |
| P13 | | 2D | - | ground/slope | - | 1 | - |
| P14 | | - | 3D | - | hill | - | 1 |
| P15 | | 2D | - | cliff surface | - | 1 | - |
| P16 | | - | 3D | - | moutainside/accessway to go up | - | 0 |
| P17 | | - | 3D | - | cliff face | - | 1 |
| P18 | | - | 3D | - | debris flow | - | 1 |
| P19 | | 2D | - | hillslope | - | 1 | - |
| P20 | | 2D | - | tree | - | 0 | - |
| P21 | | - | 3D | - | moutainside | - | 1 |
| P22 | | 2D | - | slope, man-made | - | 1 | - |
| P23 | | 2D | - | cliff face | - | 1 | - |
| P24 | | - | 3D | - | steep slope/cliff [natural] | - | 1 |
| P25 | | - | 3D | - | part of the relief/ veg on cliff | - | 1 |
| P26 | | - | 3D | - | sloping ground | - | 1 |
| P27 | | - | 3D | - | rock | - | 1 |
| P28 | | 2D | - | rocky/grassy ground | - | 1 | - |
| P29 | | 2D | - | rock or grass, part of hill | - | 1 | - |
| P30 | NA | - | - | - | - | - | - |
| P31 | | 2D | - | slope | - | 1 | - |
| P32 | | 2D | - | natural, but unsure | - | 0 | - |
| P33 | | 2D | - | very sharp slope, rocky valleys | - | 1 | - |
| P34 | | 2D | - | end of a natural slope | - | 1 | - |
| P35 | | 2D | - | tree | - | 0 | - |
| P36 | | 2D | - | don't know | - | 0 | - |
| P37 | | - | 3D | - | sloping edge | - | 1 |
| P38 | | 2D | - | natural (trees?) | - | 0 | - |
| P39 | | - | 3D | - | cliff-face | - | 1 |
| P40 | | - | 3D | - | ground | - | 1 |
| P41 | | - | 3D | - | rock, cut off for building the | - | 1 |
| P42 | | - | 3D | - | steep cliff | - | 1 |
| P43 | | 2D | - | cliff face | - | 1 | - |
| P44 | | - | 3D | - | steep drop of hillside or cliff | - | 1 |
| P45 | NA | - | - | - | - | - | - |
| P46 | | - | 3D | - | slope face | - | 1 |
| P47 | | 2D | - | ground surface on steep slope | - | 1 | - |
| P48 | | 2D | - | hillside/ground | - | 1 | - |
| P49 | NA | - | - | - | - | - | - |
| P50 | | - | 3D | - | ground worn away | - | 1 |
| P51 | | 2D | - | hill/ground | - | 1 | - |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 24 | 22 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 17 | 21 |
| total incorrect | 7 | 1 |

Scene D; POI H: tree

| particip- ant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | - | 3D | - | tree | - | 1 |
| P02 | NA | - | - | - | - | - | - |
| P03 | | 2D | - | tree | - | 1 | - |
| P04 | | 2D | - | on the ground, underneath the t | - | 0 | - |
| P05 | | - | 3D | - | part of tree | - | 1 |
| P06 | NA | - | - | - | - | - | - |
| P07 | | - | 3D | - | tree branch | - | 1 |
| P08 | | 2D | - | ground | - | 0 | - |
| P09 | | - | 3D | - | lower canopy part of tree | - | 1 |
| P10 | | 2D | - | tree | - | 1 | - |
| P11 | | - | 3D | - | tree on a hill | - | 1 |
| P12 | | 2D | - | conifer tree | - | 1 | - |
| P13 | | 2D | - | bush/tree | - | 1 | - |
| P14 | | - | 3D | - | part of tree canopy | - | 1 |
| P15 | | 2D | - | tree or rock jutting out | - | 0 | - |
| P16 | | - | 3D | - | branch or a leaf | - | 1 |
| P17 | | - | 3D | - | part of a tree | - | 1 |
| P18 | | - | 3D | - | coniferous tree | - | 1 |
| P19 | | 2D | - | shrubby | - | 1 | - |
| P20 | | 2D | - | tree | - | 1 | - |
| P21 | | - | 3D | - | bush | - | 1 |
| P22 | | 2D | - | bush | - | 1 | - |
| P23 | | 2D | - | base of tree or bush | - | 1 | - |
| P24 | | - | 3D | - | tree/vegetation | - | 1 |
| P25 | | - | 3D | - | part of a tree | - | 1 |
| P26 | | - | 3D | - | edge of tree | - | 1 |
| P27 | | - | 3D | - | tree | - | 1 |
| P28 | | 2D | - | part of a little tree | - | 1 | - |
| P29 | | 2D | - | conifer tree | - | 1 | - |
| P30 | NA | - | - | - | - | - | - |
| P31 | | 2D | - | triangular, christmas tree-shap | - | 1 | - |
| P32 | | 2D | - | tree | - | 1 | - |
| P33 | | 2D | - | vegetation | - | 1 | - |
| P34 | | 2D | - | tree | - | 1 | - |
| P35 | | 2D | - | roof | - | 0 | - |
| P36 | | 2D | - | not sure | - | 0 | - |
| P37 | | - | 3D | - | ledge or lower level undulating | - | 0 |
| P38 | | 2D | - | comms tower > maybe a tree | - | 0 | - |
| P39 | | - | 3D | - | tree | - | 1 |
| P40 | | - | 3D | - | vegetation | - | 1 |
| P41 | | - | 3D | - | tree | - | 1 |
| P42 | | - | 3D | - | lower part of conifer | - | 1 |
| P43 | | 2D | - | small tree | - | 1 | - |
| P44 | | - | 3D | - | small tree [shape + points pene | - | 1 |
| P45 | NA | - | - | - | - | - | - |
| P46 | | - | 3D | - | bush | - | 1 |
| P47 | | 2D | - | vegetation | - | 1 | - |
| P48 | | 2D | - | tree | - | 1 | - |
| P49 | NA | - | - | - | - | - | - |
| P50 | | - | 3D | - | tree | - | 1 |
| P51 | | 2D | - | tree | - | 1 | - |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 24 | 22 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 18 | 21 |
| total incorrect | 6 | 1 |

Scene D; POI I: rooftop

| particip- ant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|------------------|--------|-----|-----|---|----------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D interpretation | 3D interpretation | 2D | 3D |
| P01 | | - | 3D | - | building | - | 1 |
| P02 | NA | - | - | - | - | - | - |
| P03 | | 2D | - | Don't know | - | 0 | - |
| P04 | | 2D | - | tree crown | - | 0 | - |
| P05 | | - | 3D | - | part of a tree | - | 0 |
| P06 | NA | - | - | - | - | - | - |
| P07 | | - | 3D | - | dense area of bushes | - | 0 |
| P08 | | 2D | - | roof | - | 1 | - |
| P09 | | - | 3D | - | building roofline | - | 1 |
| P10 | | 2D | - | roof | - | 1 | - |
| P11 | | - | 3D | - | tree | - | 0 |
| P12 | | 2D | - | building | - | 1 | - |
| P13 | | 2D | - | tree | - | 0 | - |
| P14 | | - | 3D | - | canopy of the same group of tree | - | 0 |
| P15 | | 2D | - | natural mountainous surface | - | 0 | - |
| P16 | | - | 3D | - | roof | - | 1 |
| P17 | | - | 3D | - | higher up in same tree as [j] | - | 0 |
| P18 | | - | 3D | - | rock/boulder | - | 1 |
| P19 | | 2D | - | shrub or bush | - | 0 | - |
| P20 | | 2D | - | man-made structure | - | 1 | - |
| P21 | | - | 3D | - | trees | - | 0 |
| P22 | | 2D | - | roof | - | 1 | - |
| P23 | | 2D | - | tree canopy / veg | - | 0 | - |
| P24 | | - | 3D | - | roof | - | 1 |
| P25 | | - | 3D | - | top part of roof | - | 1 |
| P26 | | - | 3D | - | part of roof | - | 1 |
| P27 | | - | 3D | - | roof | - | 1 |
| P28 | | 2D | - | roof of mountain hut OR top of t | - | 0 | - |
| P29 | | 2D | - | building, house, roof | - | 1 | - |
| P30 | NA | - | - | - | - | - | - |
| P31 | | 2D | - | part of mountain (base) | - | 0 | - |
| P32 | | 2D | - | rock surface of the mountain/hi | - | 0 | - |
| P33 | | 2D | - | trees | - | 0 | - |
| P34 | | 2D | - | towards top of rocky outcrop | - | 0 | - |
| P35 | | 2D | - | roof | - | 1 | - |
| P36 | | 2D | - | something man-made | - | 1 | - |
| P37 | | - | 3D | - | ledge/slope OR undulating mount | - | 0 |
| P38 | | 2D | - | trees | - | 0 | - |
| P39 | | - | 3D | - | roof | - | 1 |
| P40 | | - | 3D | - | vegetation | - | 0 |
| P41 | | - | 3D | - | branches of tree | - | 0 |
| P42 | | - | 3D | - | tree - same as j | - | 0 |
| P43 | | 2D | - | rooftop | - | 1 | - |
| P44 | | - | 3D | - | roof of something man-made | - | 1 |
| P45 | NA | - | - | - | - | - | - |
| P46 | | - | 3D | - | top of canopy or roof | - | 0 |
| P47 | | 2D | - | vegetation | - | 0 | - |
| P48 | | 2D | - | building roof | - | 1 | - |
| P49 | NA | - | - | - | - | - | - |
| P50 | | - | 3D | - | roof | - | 1 |
| P51 | | 2D | - | roof | - | 1 | - |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 24 | 22 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 11 | 11 |
| total incorrect | 13 | 11 |

Scene D; POI J: building/rooftop

| particip- pant | method | | | interpretation of feature (abbreviated verbal answer) | | accuracy* | |
|-------------------|--------|-----|-----|---|---------------------------------|-----------|----|
| | NA? | 2D? | 3D? | 2D | 3D | 2D | 3D |
| P01 | | - | 3D | - | building | - | 1 |
| P02 | NA | - | - | - | - | - | - |
| P03 | | 2D | - | Don't know | - | 0 | - |
| P04 | | 2D | - | lower branch of canopy of veget | - | 0 | - |
| P05 | | - | 3D | - | part of a tree | - | 0 |
| P06 | NA | - | - | - | - | - | - |
| P07 | | - | 3D | - | dense area of bushes | - | 0 |
| P08 | | 2D | - | roof | - | 1 | - |
| P09 | | - | 3D | - | archway | - | 1 |
| P10 | | 2D | - | roof | - | 1 | - |
| P11 | | - | 3D | - | within the tree | - | 0 |
| P12 | | 2D | - | building | - | 1 | - |
| P13 | | 2D | - | tree | - | 0 | - |
| P14 | | - | 3D | - | canopy of the same group of tre | - | 0 |
| P15 | | 2D | - | natural moutainous surface | - | 0 | - |
| P16 | | - | 3D | - | building | - | 1 |
| P17 | | - | 3D | - | branch in a lower part of tree | - | 0 |
| P18 | | - | 3D | - | rock/boulder | - | 0 |
| P19 | | 2D | - | shrub or bush | - | 0 | - |
| P20 | | 2D | - | man-made structure | - | 1 | - |
| P21 | | - | 3D | - | trees | - | 0 |
| P22 | | 2D | - | roof | - | 1 | - |
| P23 | | 2D | - | tree canopy / veg | - | 0 | - |
| P24 | | - | 3D | - | roof | - | 1 |
| P25 | | - | 3D | - | where roof dips in, e.g. gutter | - | 1 |
| P26 | | - | 3D | - | part of roof | - | 1 |
| P27 | | - | 3D | - | roof | - | 1 |
| P28 | | 2D | - | side of moutain hut OR side of | - | 0 | - |
| P29 | | 2D | - | building, house | - | 1 | - |
| P30 | NA | - | - | - | - | - | - |
| P31 | | 2D | - | part of mountain (base) | - | 0 | - |
| P32 | | 2D | - | rock surface of the mountain/hi | - | 0 | - |
| P33 | | 2D | - | trees | - | 0 | - |
| P34 | | 2D | - | towards bottom of rocky outcrop | - | 0 | - |
| P35 | | 2D | - | roof | - | 1 | - |
| P36 | | 2D | - | tree | - | 0 | - |
| P37 | | - | 3D | - | building OR mountain | - | 0 |
| P38 | | 2D | - | trees | - | 0 | - |
| P39 | | - | 3D | - | roof | - | 1 |
| P40 | | - | 3D | - | vegetation | - | 0 |
| P41 | | - | 3D | - | branches of tree | - | 0 |
| P42 | | - | 3D | - | tree - same as i | - | 0 |
| P43 | | 2D | - | rooftop | - | 1 | - |
| P44 | | - | 3D | - | bottom edge of window or part o | - | 1 |
| P45 | NA | - | - | - | - | - | - |
| P46 | | - | 3D | - | bottom of canopy | - | 0 |
| P47 | | 2D | - | vegetation | - | 0 | - |
| P48 | | 2D | - | eaves of dormer window | - | 1 | - |
| P49 | NA | - | - | - | - | - | - |
| P50 | | - | 3D | - | part of the building | - | 1 |
| P51 | | 2D | - | border of roof | - | 1 | - |

* where 1 = correct, 0 = incorrect

sum

| | |
|----|----|
| 24 | 22 |
|----|----|

| | | |
|-----------------|----|----|
| total correct | 10 | 10 |
| total incorrect | 14 | 12 |

Examples of 2D vs. 3D verbal interpretations

Where 3D > 2D means 3D interpretation answers are significantly more accurate than 2D. 2D ≈ 3D: no significant difference in accuracy of 2D and 3D interpretation answers. N.B. there were no 2D > 3D answers, where 2D answers were significantly more accurate than 3D. The 2D and 3D quotes for each POI are unpaired – participant numbers denoted by [P. no.].

| POI | 2D example comment | 3D example comment |
|--------------------------------|--|---|
| POI B: Chimney (3D>2D) | [P40] Yeah, this is interesting because it seems a bit too far away to be part of that tree... that's there, but it doesn't see regular enough to be a chimney or something that's emerging, so whether this building has some kind of... kind of like an aerial or couple of aerials or something that is kind of an odd shape that wouldn't be represented by something that was particularly coherent. Bird? [...] | [P10] Think that one's a chimney 'cause I was underneath what I think's a roof and I looked up and there was a hole in the roof, so I went further and thought that the explanation might be because it was a chimney, because there were still points there, they just weren't at the same level as the roof. |
| POI C: hedge (2D≈3D) | [P21] This looks like this could be... yeah, I thought [previously during overview] that it could be like a covered walkway or something, but...it's quite. It's actually quite low to the ground, at this point. [...] Yeah, so... I don't- it could be like a...I still feel like it's a covered walkway or something to the, to the building. Maybe could be like a nice arch of bushes or something? Or, it could be something, like a man-made feature, but it's mostly all at the same height, so... whatever it is is probably trimmed, like, maintained. | [P31] They look like hedges, they are natural, they seem to have some little bit of pattern like flowers that have been uniformly cut or something. Yeah. They are not on the ground, they are not as high as the other trees, they look somehow little bit of shape so like flowers that had been cut. |
| POI I&J: Rooftop (2D≈3D) | [P23] The higher of the 2 [i] is vegetation in the lower of the 2 [j] is vegetation, so I'd say they're in a tree canopy. | [P44] This one here, is the roof of the structure, if it is a building. Um, actually, the roof... it looks a bit... I'm not sure if it's actually a building itself. It's something man-made because there's nothing... coming through it, so it's solid, but, it's not shaped like the rest of the roofs that we saw, it's uneven, I actually don't know what that could be, it could be... <mumble> it's too small [big?] to be a car. P44 |

Further comments on interpretation accuracy

Further comments for Q10, Scene C - How confident were you about the answers you gave when identifying the feature of the flashing points in Scene C – flat?

| P. no. | Further comment to Q10 | First Scene | Second Scene |
|--------|---|-------------|--------------|
| P40 | 1:23:19.7 Okay, the environment of C was er... I was more confident in, so maybe a... 2. I think maybe it was just more similar to the other ones we'd already looked at. | C_2D | D_3D |
| P26 | 1:40:14.5 For scene C, I was less confident. | D_3D | C_2D |
| P28 | 1:52:24.3 In the first scene, for example, the chimney looked a bit jumbled, so I wasn't quite sure [...] | C_3D | D_2D |
| P12 | 1:26:28.1 Scene C... other than that one [POI] that was in that feature where I wasn't sure if it was an attachment or something natural... | D_2D | C_3D |

Further comments for Q10, Scene D - How confident were you about the answers you gave when identifying the feature of the flashing points in Scene D – sloped?

| P. no. | Further comment to Q10 | First Scene | Second Scene |
|--------|---|-------------|--------------|
| P15 | 1:25:45.4 Well, it's harder to judge elevation, it's hard to judge because there is a natural elevation in a slope and so to notice something that is elevated even further it was harder to tell. | C_3D | D_2D |
| P34 | 1:17:11.8 As I say, with the 2D one, [I was] getting features confused between what was actually ground on the slope or raised foliage on the slope. ...(3) | C_3D | D_2D |
| P28 | 1:54:12.5 I think it's a man-made or little hut of something. because it's all jumbled on top, it doesn't fit again the picture of how a roof should look like. | C_3D | D_2D |
| P14 | 0:45:26.3 I couldn't recognise features as well, I think it was partly because the ground was sloped so couldn't tell if it was the ground going up with it was a building or tree going up. | C_2D | D_3D |
| P21 | 1:33:40.7 I wouldn't be necessarily pick out features because I don't see them everyday. [...] I don't know features that are common on steep slopes to be able to identify them in point cloud form. | C_2D | D_3D |

Appendix M: Reflection on methods (RQ3)

General 2D and 3D Observations

Participant descriptions of the 3D effect. INT denotes interviewer during dialogue.

| Scene (A/B) | Method (2D/3D) | Order (1/2) | Participant | Quote | Summary |
|-------------|----------------|-------------|-------------|--|--|
| A | 3D | 2 | P48 | <p>[Within Task] [Compared to Scene B in 2D], I'm getting more of a feel for the shape of them as well, and how they are in terms of sort of their relationship with the points around them as well?</p> <p>0:59:03.6 INT: Okay, when you say the shape of them do you mean of each point?</p> <p>0:59:08.0 Yeah, when you're looking at in 2-D, although you know that it's a round point, you're almost looking at it as a disc whereas now I feel like I'm looking at it as a ball.</p> | <p>In 3D, get a feel for relationship between points</p> <p>Not disc, more ball-shaped in 3D</p> |
| A | 2D | 1 | P37 | <p>[During Feedback] Yeah, um... I thought it [the Measurement Task] was good, I quite enjoyed it. Yes. Um... interestingly, I think I was more comfortable with the 2D, because it's more familiar. Er... yeah. But, the 3D is quite exciting because you can see more. Um, like, you feel like you're more.... involved. Does that make sense? It [3D] feels like, more real, like you're there.</p> | 2D comfortable |
| B | 3D | 2 | | | 3D exciting, "like you're there" |

Interpretation feedback - natural/man-made features

Question 9, which was posed to participants after completion of the Interpretation Task.

| | |
|----------------------|--|
| Main Question | Q9: In general, how confident were you about identifying the following types of features, in the larger scenes? |
| | Natural features, in 2D |
| | Natural features, in 3D |
| | Man-made features, in 2D |
| | Man-made features, in 3D |
| Possible Answers | 1 = Very confident; 7 = Not at all confident; 8 = don't know, 0 = n/a |

Further comments given by participants for Question 9, re. general interpretation of natural features in 3D.

| P no. | Q9 Further Comment for 'natural features in 3D' | 3D Scene Viewed |
|--------------|---|------------------------|
| P48 | 0:53:14.9 I think I was very confident, I think the trees stood out | C |
| P34 | 1:14:58.6 So, I think I'd say I'm a bit more confident with the 3D. Um...probably not VERY confident... (3) So obviously, there was what I in my initial assessment thought was some kind of walkway/decking type thing, which I think [after seeing coloured by classification] was actually a hedge, so 3 | C |
| P23 | 1:18:12.0 much more confident [than 2D] | C |
| P12 | 1:24:55.6 natural 3D – it was all pretty clear | C |
| P04 | More [confident], because you can see like the crowns of the trees. | C |
| P40 | 0:01:25.7 [d 3d] I wasn't very confident on that one, was I, so I'd just put 5 and then.... 6 | D |
| P11 | 0:42:10.2 We'll go for a 3, just because I wasn't sure all the time. | D |
| P25 | 1:34:45.1 Yeah. Um... natural features in 3D, I just think with the 3D you can get under and around, you can judge the size and the shape of the object a bit better, using the 3D. In the 2D, you kind of just sort of... more at a guess as to... what it is really. I think it [3D] helps you understand that. | D |

Further comments given by participants for Question 9, re. general interpretation of natural features in 2D.

| P no. | Q9 Further Comment for 'natural features in 2D' | 2D Scene Viewed |
|--------------|--|------------------------|
| P07 | 1:15:11.1 To be honest, it was a bit harder in 2D for the natural features, so probably say 3 for the first one again, just a bit harder to discern what was what. | C |
| P19 | 0:14:28.0 Erm, I was quite confident on the sloped one in the [natural] features, so I'd say that was about 2. | D |
| P48 | 0:52:21.3 Natural - i felt quite confident with the natural features, erm, very is maybe pushing it, so maybe a 2. | D |
| P34 | 1:12:56.3 I think the natural features in 2D, so, Scene D, um... so I was a little bit muddled with that because initially I thought it was like a rocky plane, WITHOUT any vegetation in my initial assessment, so not particularly confident in that. [...] When I was looking through with the gamepad, like, you probably noticed as I was trying to go up and down, trying to see if there were points below and things like that? Yeah, and I think that's sort of the... as I say, I THINK I found it easier to... easier to do that in the 3 dimensions in Scene C. So, to interpret the depth a little bit better, I think, to determine the points below that would suggest an irregular... um, natural feature. | D |
| P12 | 1:24:45.6 4 – middle range, so for the more obvious ones, it was alright, but the more difficult ones, it was not very clear | D |
| P04 | Natural features are harder because they've got less distinct boundaries | D |
| P36 | [P36 doesn't know is correct or not] 1:37:44.5 I'm not sure about any of them, so 4. 1:37:50.7 For all of them? Oh, you mean, you weren't sure about how you identified them? 1:37:59.1 Yeah, if I identified correctly or not | D |

Further comments given by participants for Question 9, re. general interpretation of man-made features in 2D.

| P no. | Q9 Further Comment for 'man-made features in 2D' | 2D Scene Viewed |
|--------------|---|------------------------|
| P48 | 0:52:33.5 Man-made was definitely more difficult on the slope in 2D, er... and I kind of , was looking around the feature almost for clues as to what it might be, but I still felt quite confident as to what it was when I made the guess. But, not as confident as natural features. | D |

| | | |
|-----|--|---|
| P03 | 1:20:38.6 Hmm, I don't know, I mean I didn't feel it- because there was some kind of some regular structure. For example the roofs when they were so regular, was not really much different to have 3-D or, you know, the 2-D. | D |
| P28 | 1:46:41.7 [man-made 2D] It's hard to say because that specific scene was a very jumbled roof, so I wasn't... completely confident. | D |
| P34 | 1:15:43.9 I don't think there were ANY actual features in Scene D , from my viewing of it | D |
| P19 | 0:14:07.7 Er...the man-made features in the sloped one, I don't think it's applicable, because I don't think there were any man-made features there. | D |

Further comments given by participants for Question 9, re. general interpretation of man-made features in 3D.

| P no. | Q9 Further Comment for 'man-made features in 3D' | 3D Scene Viewed |
|--------------|---|------------------------|
| P23 | 1:18:48.8 It improved both, but it was very easy, or felt very easy when their regular structures than it is when their irregular. So, so, yeah, so it was better for both but better feel for cultural structures. | C |
| P35 | 2:03:29.9 It's easier to identify man-made or natural feature in 3D than in 2D. The location of those points, the point clouds, in 3D, [are] much easier. When you're looking at the same [type of] thing in 2D, it appears [...] confusing | C |
| P40 | 0:02:08.5 [In 2D] they're [natural features] more obvious that they're not man-made because they're not dead, dead straight. I just found the whole thing less obvious in that I didn't know what it was. And because the vegetation overlaps sometimes doesn't it, the man-made, so you can't necessarily see the corners and things like that. | D |
| P50 | 1:09:21.1 I think so, yeah, because man-made things are quite... straight lines, aren't they? Or what I THOUGHT was anyway, you don't know! <laughter> | D |
| P25 | 1:35:12.0 It seems a lot easier [in both 2D and 3D for man-made features], just because, er, the linearity of it. It looks like it's a bit more uniform and because the points tend to run in patterns, you can understand... once you get an understanding of a pattern, you can use that pattern to then replicate your knowledge of an object of shape to what you think an object could be. | D |

Appendix N: Digital appendices

Digital data can be found on the CD affixed to back cover. Files include:

README file

Further information regarding digital files.

.txt files - AOIs for scenes A, B, C, and D.

Lidar data © Airbus Defence and Space Ltd. (2013a and 2013b). See Chapter 3 for full data processing.

.py files - Visualisation modules

For use with Vizard 3.0 (Worldviz, 2010b) or later versions. Visualisation system configurations are detailed in Chapter 3.

