Natural Freehand Grasping of Virtual Objects for Augmented Reality

Maadh Al Kalbani

A thesis presented in partial fulfilment for the degree of Doctor of Philosophy

Digital Media Technology Lab

Birmingham City University

United Kingdom

September 2018

Abstract

Grasping is a primary form of interaction with the surrounding world, and is an intuitive interaction technique by nature due to the highly complex structure of the human hand. Translating this versatile interaction technique to Augmented Reality (AR) can provide interaction designers with more opportunities to implement more intuitive and realistic AR applications. The work presented in this thesis uses quantifiable measures to evaluate the accuracy and usability of natural grasping of virtual objects in AR environments, and presents methods for improving this natural form of interaction.

Following a review of physical grasping parameters and current methods of mediating grasping interactions in AR, a comprehensive analysis of natural freehand grasping of virtual objects in AR is presented to assess the accuracy, usability and transferability of this natural form of grasping to AR environments. The analysis is presented in four independent user studies (120 participants, 30 participants for each study and 5760 grasping tasks in total), where natural freehand grasping performance is assessed for a range of virtual object sizes, positions and types in terms of accuracy of grasping, task completion time and overall system usability.

Findings from the first user study in this work highlighted two key problems for natural grasping in AR; namely inaccurate depth estimation and inaccurate size estimation of virtual objects. Following the quantification of these errors, three different methods for mitigating user errors and assisting users during natural grasping were presented and analysed; namely dual view visual feedback, drop shadows and additional visual feedback when adding user based tolerances during interaction tasks. Dual view visual feedback was found to significantly improve user depth estimation, however this method also significantly increased task completion time. Drop shadows provided an alternative, and a more usable solution, to dual view visual feedback through significantly improving depth estimation, task completion time and the overall usability of natural grasping. User based tolerances negated the fundamental problem of inaccurate size estimation of virtual objects, through enabling users to perform natural grasping without the need of being highly accurate in their grasping performance, thus providing evidence that natural grasping can be usable in task based AR environments.

Finally recommendations for allowing and further improving natural grasping interaction in AR environments are provided, along with guidelines for translating this form of natural grasping to other AR environments and user interfaces.

Acknowledgements

I would firstly like to thank my family and especially my parents for their support and patience over the years:

إهـداء إلـى أبـي و أمـي العزيـزين، أشـكركم جزيـل الشـكر علـى دعمكـم اللامحـدود خـلال فتـرة دراسـتي و غربتـي، وثقـتكم بقـدرتي علـى إجتيـاز هـذه المرحلـة الشـاقة بنجـاح. كـان وسـيبقى أملـي دائمـاً أن تفخـروا بـي، وأن يقـدرني الله علـى أن أرد كـل مـا قـدمتوه لأجلـي من مساعدات و تضحيات و صبر. كنتم بجانبي في كل خطوة في هذا الطريق، وستكونون دائماً رفقاء كل طريق.

Thanks to all members of the DMT Lab that helped in one way or another in this work. Especially my friends (Sean Enderby, Sam Smith, Matthew Cheshire, Dalia El Banna, Alan Dolhasz, Greg Hough, Nick Jillings, Carl Southall) that were more like a family to me and helped me through tough times, and were a great source of inspiration, laughter and pure joy.

To my supervisors, Dr. Ian Williams, Prof. Cham Athwal and Dr. Maite Frutos, I honestly can't thank you enough for your support, patience and non-stop feedback on my work. This would have never been accomplished without your guidance and support. I was once told by the head of research at BCU that "I've had the best research training in the DMT Lab", and this says it all about this supervision team. Special thanks to Ian for believing in me, and for all the support that he offered both on a personal and professional level even when he didn't have to. It really made a huge difference.

I would also like to thank the Omani Ministry of Higher Education, that never stopped supporting me in order to achieve this dream of mine. I will always be indebted to their generosity and support over the years.

Special thanks to Janet Hunt, for saving my life at some point during this PhD. I would have never been able to make it to this stage without you sticking by side when I needed it the most. I will forever be grateful, and I am more than honoured to call you a friend.

To Moosa Al Shuraiqi, Mohammed Al Rahbi, Ahmed Al Ghafri and Khalid Al Kharusi, we've rarely had a chance to be co-located for long, but you always found the time to be by my side when needed even if took crossing all sorts of geographical boundaries. Thanks for all the support

Last but definitely not least, my beautiful wife, and the real MVP, Dr. Reham Ismail. I can't thank you enough for your endless and unconditional support, and for all the sacrifices you made in this long bumpy way. You made this journey much easier, and your smile always kept my sanity intact. Let's keep defying odds :)

Contents

1 Introduction **1**

1.1 Motivation . . . 1
1.2 Research Objectives . . . 3
1.3 Thesis Structure . . . 4
1.4 Contributions . . . 6
1.5 Published Papers . . . 6

2 Hand Based Interaction in Mixed and Augmented Reality **8**

2.1 Introduction . . . 8
2.2 Virtuality Definitions . . . 9
2.3 Virtual Object Interaction . . . 11
2.3.1 Interaction Definition . . . 11
2.3.2 Interaction Framework . . . 11
2.3.3 User Interfaces . . . 12
2.3.4 Interaction Techniques . . . 13
2.4 Natural Interaction . . . 14
2.4.1 Grasping Techniques . . . 14
2.4.1.1 Wearable Based . . . 15
2.4.1.2 Freehand Based . . . 17
2.5 Freehand Grasping of Virtual Objects . . . 22
2.5.1 Definition . . . 23
2.5.2 Rationale . . . 24
2.5.3 Applications . . . 25
2.5.4 Problems . . . 26
2.6 Summary . . . 27

3 Grasping Interaction **29**

3.1 Introduction . . . 29
3.2 Grasping Definition . . . 30
3.3 Grasping Physical Objects . . . 30
3.3.1 Biomechanics of Grasping . . . 30

3.3.2 Neurophysiology of Grasping . . . 32
3.4 Grasp Classification . . . 34
3.4.1 Descriptive Taxonomies . . . 34
3.4.2 Objective Quantification . . . 40
3.5 Grasp Planning . . . 41
3.5.1 Grasp Constraints . . . 41
3.5.2 Grasp Phases . . . 43
3.5.2.1 Planning . . . 44
3.5.2.2 Reaching . . . 45
3.5.2.3 Pre-load . . . 45
3.5.2.4 Load . . . 45
3.5.2.5 Transition . . . 45
3.5.2.6 Release . . . 46
3.6 Summary . . . 46

4 Freehand Grasp Evaluation Methodology **47**
4.1 Introduction . . . 47
4.2 Methodology . . . 48
4.2.1 Grasp Parameters . . . 48
4.2.1.1 Grasp Taxonomy . . . 48
4.2.1.2 Grasp Type . . . 49
4.2.1.3 Grasp Phases . . . 50
4.2.2 Grasp Metrics . . . 50
4.2.2.1 Grasp Aperture . . . 50
4.2.2.2 Grasp Displacement . . . 51
4.2.3 Baseline Environment . . . 54
4.2.3.1 Setup Overview . . . 54
4.2.3.2 System Architecture . . . 56
4.2.3.3 Tracking . . . 58
4.2.3.4 Virtual Objects . . . 58
4.2.3.5 Feedback . . . 60
4.2.4 Experiment Protocol . . . 62
4.2.4.1 Participants . . . 62
4.2.4.2 Protocol . . . 62
4.3 Studies and Hypotheses . . . 62
4.4 Summary . . . 65

5 Study 1: Freehand Grasping Errors **66**
5.1 Introduction . . . 66
5.2 Study Outline . . . 67

5.2.1 Design . . . 67
5.2.2 Participants . . . 68
5.2.3 Statistical Model . . . 68
5.2.4 Protocol . . . 68
5.3 Experiment 1 - Object Size . . . 69
5.3.1 Experiment Design . . . 69
5.3.2 Procedure . . . 69
5.3.3 Results . . . 69
5.3.3.1 Results - Grasp Aperture (GAp) . . . 70
5.3.3.2 Analysis - Grasp Aperture (GAp) . . . 70
5.3.3.3 Results - Completion Time . . . 71
5.3.3.4 Analysis - Completion Time . . . 72
5.3.3.5 Results - Grasp Displacement ($GDisp$) . . . 72
5.3.3.6 Analysis - Grasp Displacement ($GDisp$) . . . 72
5.3.3.7 Results - Object Type . . . 76
5.3.3.8 Analysis - Object Type . . . 76
5.4 Experiment 2 - Object Position . . . 77
5.4.1 Experiment Design . . . 77
5.4.2 Procedure . . . 77
5.4.3 Results . . . 78
5.4.3.1 Results - Grasp Aperture (GAp) . . . 78
5.4.3.2 Analysis - Grasp Aperture (GAp) . . . 78
5.4.3.3 Results - Completion Time . . . 82
5.4.3.4 Analysis - Completion Time . . . 84
5.4.3.5 Results - Grasp Displacement ($GDisp$) . . . 85
5.4.3.6 Analysis - Grasp Displacement ($GDisp$) . . . 85
5.4.3.7 Results - Object Type . . . 93
5.4.3.8 Analysis - Object Type . . . 93
5.5 Conclusions . . . 96

6 Study 2: Dual View Visual Feedback for Freehand Grasping 98
6.1 Introduction . . . 98
6.2 Study Outline . . . 99
6.2.1 Design . . . 99
6.2.2 Participants . . . 102
6.2.3 Statistical Analysis . . . 103
6.2.4 Protocol . . . 103
6.3 Experiment 1 - Object Size . . . 104
6.3.1 Experiment Design . . . 104

6.3.2 Procedure . . . 104
6.3.3 Results . . . 104
6.3.3.1 Results - Grasp Aperture (GAp) . . . 104
6.3.3.2 Analysis - Grasp Aperture (GAp) . . . 104
6.3.3.3 Results - Completion Time . . . 106
6.3.3.4 Analysis - Completion Time . . . 106
6.3.3.5 Results - Grasp Displacement ($GDisp$) . . . 106
6.3.3.6 Analysis - Grasp Displacement ($GDisp$) . . . 107
6.3.3.7 Results - Object Type . . . 111
6.3.3.8 Analysis - Object Type . . . 111
6.3.3.9 Usability Analysis . . . 112
6.4 Experiment 2 - Object Position . . . 113
6.4.1 Experiment Design . . . 113
6.4.2 Procedure . . . 113
6.4.3 Results . . . 114
6.4.3.1 Results - Grasp Aperture (GAp) . . . 114
6.4.3.2 Analysis - Grasp Aperture (GAp) . . . 114
6.4.3.3 Results - Completion Time . . . 120
6.4.3.4 Analysis - Completion Time . . . 120
6.4.3.5 Results - Grasp Displacement ($GDisp$) . . . 122
6.4.3.6 Analysis - Grasp Displacement ($GDisp$) . . . 122
6.4.3.7 Results - Object Type . . . 131
6.4.3.8 Analysis - Object Type . . . 132
6.4.3.9 Usability Analysis . . . 134
6.5 Conclusions . . . 135

7 Study 3: Drop Shadows for Freehand Grasping 138
7.1 Introduction . . . 138
7.1.1 Drop Shadows as a Depth Cue . . . 138
7.2 Study Outline . . . 139
7.2.1 Design . . . 139
7.2.2 Participants . . . 142
7.2.3 Statistical Analysis . . . 142
7.2.4 Protocol . . . 142
7.2.5 Procedure . . . 143
7.3 Results . . . 143
7.3.1 Results - Grasp Aperture (GAp) . . . 143
7.3.2 Analysis - Grasp Aperture (GAp) . . . 144
7.3.3 Results - Completion Time . . . 150

7.3.4 Analysis - Completion Time . . . 150
7.3.5 Results - Grasp Displacement ($GDisp$) . . . 153
7.3.6 Analysis - Grasp Displacement ($GDisp$) . . . 154
7.3.7 Results - Object Type . . . 163
7.3.8 Analysis - Object Type . . . 165
7.3.9 Usability Analysis . . . 168
7.4 Conclusions . . . 169

8 Study 4: User Based Tolerances for Freehand Grasping 172
8.1 Introduction . . . 172
8.1.1 Interaction Tolerances . . . 173
8.2 Study Outline . . . 174
8.2.1 Interaction Grasping Tasks . . . 174
8.2.2 Interaction Offsets . . . 175
8.2.3 Design . . . 177
8.2.4 Participants . . . 179
8.2.5 Statistical Analysis . . . 179
8.2.6 Protocol . . . 179
8.2.7 Procedure . . . 180
8.3 Results . . . 181
8.3.1 Results - Completion Time . . . 181
8.3.2 Analysis - Completion Time . . . 181
8.3.3 Usability Analysis . . . 181
8.4 Conclusions . . . 185

9 Discussion and Recommendations 187
9.1 Grasp Aperture (GAp) . . . 188
9.2 Grasp Displacement ($GDisp$) . . . 192
9.3 Usability and Design Recommendations . . . 199
9.3.1 Virtual Object Size . . . 200
9.3.2 Grasp Type . . . 200
9.3.3 Grasp Phases . . . 201
9.4 Summary . . . 203

10 Conclusions 204
10.1 Review of Research . . . 206
10.1.1 Measuring Accuracy of Freehand Grasping in AR . . . 206
10.1.2 User Studies . . . 206
10.2 Constraints and Future Work . . . 208
10.2.1 Environment . . . 208

10.2.2 Methods and Transferability 209
10.2.3 Interaction Technique 210

A Post-Test Questionnaires - Dual View Visual Feedback (Chapter 6) 211

B Post-Test Questionnaires - Drop Shadows (Chapter 7) 216

C Post-Test Questionnaires - User Based Tolerances (Chapter 8) 222

List of Figures

1.1 Framework of the four user studies presented in this thesis 4

2.1 Virtuality continuum presented by Milgram and Colquhoun (1999). The red circle indicates the position of this work within this continuum. Image adapted from Milgram and Colquhoun (1999) 9

2.2 Examples of wearable based interactions in AR environments 16

2.3 Examples of virtual grasping and mobile wearable interaction approaches in AR environments 17

2.4 Examples of early work in tabletop freehand interaction techniques with virtual objects projected on a surface. Images courtesy of Rekimoto (2002); Wu and Balakrishnan (2003) 18

2.5 Examples of early work addressing occlusion problems and hand gestures (including grasping) in AR. Images courtesy of Lee and Kim (2004); Buchmann et al. (2004) 18

2.6 Examples of recent tabletop freehand interaction systems 19

2.7 Examples of handheld AR that facilitate freehand grasping interactions 20

2.8 Examples of freehand interactions and manipulation of virtual menus 21

2.9 Examples of freehand interaction in AR environments without the use of additional sensors on the hand or arm 22

2.10 Examples of freehand grasping applications in AR environments 26

3.1 Human anatomy of the hand and wrist. The human hand consists of 8 carpal bones in the wrist, 5 metacarpal bones in the palm, 2 phalanges in the thumb and 3 phalanges in each of the four fingers 31

3.2 CNS areas involved in planning, programming and execution of grasping movements. First: the cerebral cortex, basal ganglia and cerebellum plan the grasping action prior to any movement. Second: the association cortex, limbic cortex, lateral cerebellum and basal ganglia send the intended/planned motor commands to the motor cortex. Third: the motor cortex then executes movement (trajectory planning). Fourth: sensory information is then fed back to cortical areas for further planning (e.g. grasp corrections). Image courtesy of MacKenzie and Iberall (1994) 32

3.3 Areas of the CNS that are involved in grasping movements. Image courtesy of Kelly et al. (1985) 34

3.4 First grasp taxonomy by Schlesinger (1919) that treated the hand as a tool, and took into account object shape (cylindrical or spherical), hand shape (hook, open fist, close fist), and hand surfaces (tip, palmar, lateral). Image courtesy of Schwarz and Taylor (1955) 35

3.5 Napier (1956) presented a taxonomy that focused on task requirements (power or precision), and argued that the hand could form either a power or precision grasp to match any task requirements. Images courtesy of Napier (1956) 36

3.6 Examples of dynamic grasps: Landsmeer (1962) revised the definition of a precision and highlighted the importance of the interplay between the index, thumb and middle finger in the human hand. This lead to presenting new grasp types, that are variations of this interplay 36

3.7 Full descriptive taxonomies in current research for physical grasping 37

3.8 The three ways a hand provides opposition around objects, where combinations of these ways form different grasp postures. Solid black lines show the opposition vectors in the object, and the shaded area shows the plane of the palm. Images courtesy of MacKenzie and Iberall (1994) 38

3.9 Oppositions described in terms of virtual fingers. Direction of virtual fingers is parallel to the plam in the Pad opposition, perpendicular to the plam in the Palm opposition and transverse to the palm in the Side opposition. Image courtesy of MacKenzie and Iberall (1994) 38

3.10 The most versatile grasp types for the majority of graspable objects. A higher grasp span means a higher versatility for a given grasp type. Image courtesy of Bullock et al. (2013) 39

3.11 The three main grasp constraints that should be met in order to perform or choose an optimal grasp type. Image courtesy of León et al. (2014) 42

3.12 Grasp phases: a grasp starts from a resting posture in the pre motor planning stage (Opposition Space Planning), followed by opening of the hand in preparation for contact with the object (Setting Up Opposition Space). Manipulation or translation of the object then occurs once contact is made with the object (Using Opposition Space), and then finally the opposition space (i.e. object) is released. During these phases, motor commands from the CNS are generated at the Opposition Space, Biomechanical and Sensorimotor levels for the movement to satisfy all the constraints in a given task. Image courtesy of MacKenzie and Iberall (1994) 44

4.1 Full Hand Wrap Grasps 49

4.2 4.2a Grasp type (Medium Wrap) analysed in this work. 4.2b Grasp Aperture (GAp) used for quantifying grasp accuracy 50

4.3 4.3a Overestimation and 4.3b underestimation of virtual objects size measured using GAp 51

4.4 gmp that is used for measuring Grasp Displacement ($GDisp$) in the x ($GDisp_{\text{x}}$), y ($GDisp_{\text{y}}$) and z ($GDisp_{\text{z}}$) axes 52

4.5 Direction of gmp placement in relation to the omp in the x (4.5a and 4.5b) and y (4.5c and 4.5d) axes measured using $GDisp_{\text{x}}$ and $GDisp_{\text{y}}$ 53

4.6 4.6a Overestimation and 4.6b underestimation of the virtual object's position in the z axis measured using $GDisp_{\text{z}}$ 54

4.7 Setup of the system developed 55

4.8 Virtual objects in this work 59

4.9 Visual feedback in exocentric AR 60

4.10 Occlusion handling in this work 61

4.11 Structure of the four studies in this work 64

5.1 Users were able to fully occlude objects that were 40mm in size. This size showed the lowest accuracy in terms of matching GAp to object size 71

5.2 GAp for different object sizes in the 1600mm z plane. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range . . 71

5.3 gmp placement in the x and y axes of all participants in six virtual object sizes (in the order 40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 5.3a: black squares indicate cube sizes. 5.3b: black circles indicate sphere sizes. Density heat maps indicate gmp placement across participants (red indicates higher density) 73

5.4 gmp placement in the z axis of all participants in six virtual object sizes (in the order 40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 5.4a: black squares indicate cube sizes. 5.4b: black circles indicate sphere sizes. Density heat maps indicate gmp placement across participants (red indicates higher density) 75

5.5 GAp for different object positions in the three z planes in this experiment. 5.6c: 1400mm z plane. 5.6b: 1600mm z plane. 5.6a: 1800mm z plane. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range 80

5.6 Users showed higher accuracy in terms of matching GAp to objects size in the middle 1600mm z plane (5.6b) due to the spatial convenience of this position for users 82

5.7 Bottom positions were problematic for users in size estimation, especially in the furthest z plane from users where leaning forward was required to correctly grasp the object 82

5.8 Optimal interaction regions for users across all z planes in terms of all the measurement used in this work to assess grasp accuracy. X axis refers to Left, Centre and Right positions. Y axis refers to Top, Centre and Bottom positions. Values presented are Means $\pm$ SD of each corresponding measurement. Most accurate/quickest region is marked with a star 83

5.9 gmp placement in the x and y axes of all participants in 27 positions in 3 z planes (in the order: 1400mm z plane - 1600mm z plane - 1800mm z plane). 5.9a: black squares indicate cube sizes. 5.9b: black circles indicate sphere sizes. Density heat maps indicate gmp placement across participants (red indicates higher density) . 87

5.10 gmp placement in the z axis of all participants for gras[ing cubes in 27 positions in 3 z planes (First row: 1400mm z plane. Second row: 1600mm z plane. Third row: 1800mm z plane): squares indicate cube positions (Left, Centre and Right), and density heat maps indicate gmp placement across participants (red indicates higher density) 90

5.11 gmp placement in the z axis of all participants for grasping spheres in 27 positions in 3 z planes (First row: 1400mm z plane. Second row: 1600mm z plane. Third row: 1800mm z plane): squares indicate sphere positions (Left, Centre and Right), and density heat maps indicate gmp placement across participants (red indicates higher density) 91

5.12 Top positions are prone to the forearm intersecting the FOV of users in the current setting, this inability to clearly visualise the full hand can hinder depth estimation 93

5.13 An example showing how the edges of cubes, even if virtual, can sometimes dictate the grasp structure and task completion time as users try to perform grasps that comply with the geometrical structure of the cubes presented 95

6.1 Setup of the dual view visual feedback system developed 101

6.2 GAp for different object sizes in the 1600mm z plane using: 6.2a [page 106] Single view visual feedback and 6.2b [page 106] Dual view visual visual feedback. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range 106

6.3 gmp placement in the x and y axes for cubes and spheres of all participants in six sizes (40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 6.3a and 6.3c: Single view visual feedback. 6.3b and 6.3d: Dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density) . 108

6.4 gmp placement in the z axis for cubes and spheres of all participants in six sizes (40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 6.4a and 6.4c: Single view visual feedback. 6.4b and 6.4d: Dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density) 110

6.5 GAp for different object positions in the three z planes in this experiment (1400mm, 1600mm and 1800mm). 6.5a: Single view visual feedback. 6.5b: Dual view visual feedback. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range 116

6.6 Users showed higher accuracy in terms of matching GAp to objects size in the closest 1800mm z plane. This is potentially attributed to the placement of the object in relation to the body of the user, where the body in this particular plane acts as a spatial cue for the virtual object presented 117

6.7 Optimal interaction regions for users across all z planes in terms of all the measurement used in this work to assess grasp accuracy. X axis refers to Left, Centre and Right positions. Y axis refers to Top, Centre and Bottom positions. Values presented are Means $\pm$ SD of each corresponding measurement. Most accurate/quickest region is marked with a star 119

6.8 Users consistently showed the lowest accuracy in terms of matching GAp to objects size in Right positions. This is potentially attributed to the position of the side view window on the feedback monitor that is on the right hand side of users, and this lead users to move their upper body or just their heads in order to be able to correct their GAp 120

6.9 gmp placement in the x and y axes for cubes and spheres of all participants in 3 z planes (1400mm - 1600mm - 1800mm). 6.9a and 6.9c: Single view visual feedback. 6.9b and 6.9d: Dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density) 125

6.10 gmp placement in the z axis for cubes (black squares) and spheres (black circles) of all participants in 3 z planes (starting from top row in the order: 1400mm - 1600mm - 1800mm) using 6.10a: single view visual feedback and 6.10b: dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density) 130

7.1 Setup of the dual view visual feedback system developed 140
7.2 GAp for different object positions in the three z planes in this experiment (2000mm, 2200mm and 2400mm). 7.2a: Drop shadows used. 7.2b: No drop shadows used. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range 146
7.3 **Drop Shadows Condition:** Optimal interaction regions for users across all z planes in all the measurement used in this work to assess grasp accuracy. X axis refers to Left - Centre - Right positions. Y axis refers to Top - Centre - Bottom positions. Values presented are Means ± SD of each measurement. Most accurate/quickest region is marked with a star 148
7.4 **No Drop Shadows Condition:** Optimal interaction regions for users across all z planes in all the measurement used in this work to assess grasp accuracy. X axis refers to Left - Centre - Right positions. Y axis refers to Top - Centre - Bottom positions. Values presented are Means ± SD of each measurement. Most accurate/quickest region is marked with a star 149
7.5 Task completion times for different object positions in the three z planes in this study (2000mm, 2200mm and 2400mm). 7.5a: Cubes. 7.5b: Spheres. White points on boxplots indicate the mean completion time across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range 152
7.6 gmp placement in the x and y axes for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm). 7.6a and 7.6c: Drop shadows used. 7.6b and 7.6d: No Drop shadows used. Density heat maps indicate gmp placement across participants (red indicates higher density) 156
7.7 gmp placement in the z axis for cubes (black squares) and spheres (black circles) of all participants in 3 z planes (starting from top row in the order: 1400mm - 1600mm - 1800mm) using 7.7a: drop shadows and 7.7b: No drop shadows. Density heat maps indicate gmp placement across participants (red indicates higher density) 162

8.1 The two types of user based tolerances assessed in this Study. 8.1a: Absolute tolerances that are position based (unique to each position), and 8.1b: that are area based (per each z plane) 176

8.2 The 4 tasks that participants completed in this study. Letter A: starting location of the virtual objects. Letter B: target location. Arrows: motion direction. Distance between starting and target locations was constantly 400mm. Results in this study are only analysed in the reaching and pre-load phases of interaction (i.e. point A) 177
8.3 Exocentric AR system in this study. 8.3a: experiment setup, where X marks the standing position of participants. 8.3b: Feedback presented to participants during reach to grasp tasks. A and B yellow circles are the starting (A) and target (B) locations and the red and green circles show the state of the grasp. Top: participant locating the object (circle remains red). Bottom: participant successfully grasped the object (circle turns green) 178
8.4 Task completion times for the reaching and pre-load phases of interaction in the four tasks in the three z planes in this study (1400mm, 1600mm and 1800mm). 8.4a: Cubes. 8.4b: Spheres. White points on boxplots indicate the mean completion time across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range 183

9.1 Measuring user grasping accuracy using GAp in 9.1a: this work, and 9.1b: previous research 188
9.2 Visual feedback methods used in this work 9.2a: single view visual feedback using a single monitor, 9.2b: dual view visual feedback using an additional side camera and 9.2c: drop shadows alongside occlusion 189
9.3 Adaptive methods for assisting in freehand interaction with virtual objects in AR in 9.3a: this work, and 9.3b: previous research 191
9.4 Distance from target metrics used to assess user accuracy in interaction and grasping in previous research. From left to right: images courtesy of: Swan et al. (2015), Chen and Saunders (2016) and Kim and Park (2015) 193
9.5 Perceptual vergence - accommodation mismatches in 9.5a: this work, and 9.5b: previous research in AR 194
9.6 Example of improved depth estimation using dual view visual feedback 195
9.7 Drop shadows assessment in 9.7a: this work, and 9.7b: previous research in AR . . 196
9.8 User based tolerances aided users in performing natural grasping without requiring them to be highly accurate in estimating the positions and sizes of virtual objects 198
9.9 Arm reach assessment on interaction accuracy in 9.9a: this work, and 9.9b: previous research 199
9.10 Presenting users with small virtual objects can hinder usability, and can lead to users to 9.10a: fully occlude virtual objects or 9.10b: change their grasp type and posture to show parts of the virtual object during freehand grasping 200

List of Tables

3.1 Summary of the different types of grasp constraints. Table adapted from Iberall and MacKenzie (1990) 43

5.1 Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor 67

5.2 Significant Post-hoc Pairwise Comparisons - Experiment 1. Symbols are represented in constant order (● ○ ▲ △ ◆ ◇ ■ □) and represent significance in a post-hoc Dunn Test with Bonferroni correction using an α level of 0.01 for the following: ● GAp - Cube, ○ GAp - Sphere, ▲ $GDisp_{\mathrm{x}}$ - Cube, △ $GDisp_{\mathrm{x}}$ - Sphere, ◆ $GDisp_{\mathrm{y}}$ - Cube, ◇ $GDisp_{\mathrm{y}}$ - Sphere, ■ $GDisp_{\mathrm{z}}$ - Cube, □ $GDisp_{\mathrm{z}}$ - Sphere. No symbols indicate statistical similarity 69

5.3 Descriptive Statistics of GAp, $GDisp_{\mathrm{x}}$, $GDisp_{\mathrm{y}}$, $GDisp_{\mathrm{z}}$ and completion time for different sizes of cubes and spheres (Mean ± SD). For statistical significance (p < 0.01) between individual sizes, see Table 5.2 [page 69] 70

5.4 Significant Post-hoc Pairwise Comparisons - Experiment 2. Symbols are represented in constant order (● ○ ▲ △ ◆ ◇ ■ □) and represent significance in a post-hoc Dunn Test with Bonferroni correction using an α level of 0.01 for the following: ● GAp - Cube, ○ GAp - Sphere, ▲ $GDisp_{\mathrm{x}}$ - Cube, △ $GDisp_{\mathrm{x}}$ - Sphere, ◆ $GDisp_{\mathrm{y}}$ - Cube, ◇ $GDisp_{\mathrm{y}}$ - Sphere, ■ $GDisp_{\mathrm{z}}$ - Cube, □ $GDisp_{\mathrm{z}}$ - Sphere. No symbols indicate statistical similarity 79

5.5 Descriptive Statistics of GAp for different positions of cubes and spheres (Mean ± SD). For statistical significance (p < 0.01) between individual positions in each z plane, see Table 5.4 [page 79] 81

5.6 Descriptive Statistics of Task Completion Time (Mean ± SD) 84

5.7 Descriptive Statistics of $GDisp_{\mathrm{x}}$ (Mean ± SD) 86

5.8 Descriptive Statistics of $GDisp_{\mathrm{y}}$ (Mean ± SD) 88

5.9 Descriptive Statistics of $GDisp_{\mathrm{z}}$ (Mean ± SD) 92

5.10 Descriptive Statistics of Object Type (Mean ± SD). Significant differences between cubes and spheres (p < 0.01) are marked with (*) 94

6.1 Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor 102

6.2 Descriptive Statistics of GAp, $GDisp_x$, $GDisp_y$, $GDisp_z$ and completion time for different sizes of cubes and spheres (Mean ± SD). Statistical significance ($p < 0.01$) between between single and dual view visual feedback methods are marked with (*) 105

6.3 Descriptive Statistics of GAp (Mean ± SD. Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*) 115

6.4 Descriptive Statistics of Task Completion Time (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*) 121

6.5 Descriptive Statistics of $GDisp_x$ (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*) 123

6.6 Descriptive Statistics of $GDisp_y$ (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*) 127

6.7 Descriptive Statistics of $GDisp_z$ (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*) 129

6.8 Descriptive Statistics of Object Type (Mean ± SD). Significant differences ($p < 0.01$) between cubes and spheres in the dual view visual feedback condition are marked with (*) 133

7.1 Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor 141

7.2 Descriptive Statistics of GAp (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*) . . 145

7.3 Descriptive Statistics of Task Completion Time (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*) 151

7.4 Descriptive Statistics of $GDisp_x$ (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*) . . 155

7.5 Descriptive Statistics of $GDisp_y$ (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*) . . 158

7.6 Descriptive Statistics of $GDisp_z$ (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*) . . 161

7.7 Descriptive Statistics of Object Type (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*) 166

8.1 Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor 179

8.2 Descriptive Statistics of Task Completion Time (Mean ± SD). Significant differences (p < 0.01) between absolute and average user based grasp tolerances are marked with (*) 182

10.1 Summary of findings across the four user studies in this thesis. Methods that had a significant impact in improving freehand grasping in terms of the four measures used in this work (GAp, $GDisp_x$, $GDisp_y$, $GDisp_z$, task completion time) are marked with (✓). Methods that resulted in comparable results between conditions or had no significant impact on grasping accuracy and usability are marked with (–) 205

10.2 Summary of usability findings across the three studies in this thesis where usability was measured (1-3). Scores and labels are based on the SUS usability test (Brooke, 1996; Bangor et al., 2009*a*) 206

Chapter 1

Introduction

1.1 Motivation

Imagine you are a construction engineer in an Augmented Reality (AR) system, where your view is augmented with virtual objects in a real environment in real-time. The aim of this AR system you are integrated in is to train you using the same tools that you normally use in the real working environment to improve your skills in construction engineering in a risk free environment. Accordingly you are presented with two co-located tools for a given task, where one is real and another is virtual, rendered in high visual quality. You are now required to interact with both tools using your hands to complete the training task, thus you naturally grasp the real tool as that is how you would normally interact with it in real life. For the virtual tool, the system you are in is highly realistic and you can clearly visualise the dimensions and position of the virtual tool in three-dimensional space. You try and grasp the virtual tool as you did with the real tool, however the system you are in does not allow you to grasp it, but instead you are to use a different mid-air gesture-based interaction technique for interacting with virtual objects (i.e. air tap). Therefore two different interactions are required for real and virtual objects, this presents the following questions: how can interacting with virtual objects improve skills if I am using a different interaction method than the one I normally use in the real environment? If I am a novice user in AR environments that I will most likely only use for training purposes, should not the system I am using for training allow me to use a natural form of interaction that I am already familiar with given my experience in real environments? Why should one familiarise themselves with an interaction technique that is not naturally used in real environments when the virtual object could be interacted with using a natural form of interaction? Would not one be more accurate in interaction with virtual objects if a natural interaction technique is used? and how would one know when to use a natural form of interaction or a different gesture interaction technique when presented with virtual objects? These questions are particularly important when AR technologies are used as training tools for real life applications. For example, engineering and manufacturing assembly tasks that use AR technologies aim to improve the knowledge

and expertise of users in task completion and performance, however recent research shows that current interaction techniques used by current AR devices do not provide the required accuracy and usability to be truly effective for real life applications (Evans et al., 2017). This raises the final question of: could users be better in terms of performance in their real life tasks if they are using different interaction techniques in the AR environments that they are trained in? and could AR offer a truly powerful training system if more natural interactions were employed throughout?

Interaction with virtual objects can be achieved using a variety of innovative interaction techniques with the use of dexterity and coordination of body parts such as eyes, head and hands for interaction in AR now more common due to technological advances in motion tracking. Current research offers various interaction techniques that range from basic gesture and touch-based interaction techniques to the more complex brain-computer interaction. Commonly, the human hand is widely used as an interaction tool in AR environments, mainly due to its natural use on a daily basis in real life and its various skeletal and muscular degrees of freedom that allow extremely dexterous interactions and postures. Current research offers evidence for the wide use, benefits and potential problems of hand based virtual object interaction. Analysis of the accuracy and usability of the most complex subset of hand based virtual object interaction, that is grasping, in AR environments is still largely unexplored. Grasping as an interaction technique in AR systems is currently presented with a clear detachment from the concepts and parameters of the extensively researched topic of physical grasping, thus how well physical grasping performs in AR in terms of accuracy and usability is currently unclear. Current research in AR presents grasping and other interaction techniques that are primarily dictated by the software and hardware capabilities used, this leads users to perform interaction techniques that are not normally used in real environments (e.g. in the work of Cidota et al. (2016) users performed unnatural grasping of virtual objects in order to ensure their hand is oriented towards the wearable depth sensor used to asses upper motor impairments). In addition, grasping in current literature largely ignores grasp parameters that can potentially have a significant impact on interaction accuracy and usability such as task constraints, grasp type, hand dexterity, grasp phases and grasp spatial positioning in relation to objects. More advanced AR systems present grasping that is supported by additional hand sensors in order to recreate the haptic feedback experienced in physical grasping (e.g. Pacchierotti et al. (2012); Sutherland et al. (2013); Vieira et al. (2015)). However, this approach is prone to fatigue and discomfort for users due to the cumbersome additional sensors, and the accuracy of grasping using such settings is also not addressed. Moreover, additional sensors are not practical or feasible to use in certain domains such as medical applications. For medical applications current research uses AR technologies to assist users through improved feedback methods (e.g. Inoue et al. (2013)) and visualisations (e.g. Nicolau et al. (2011)), and avoid disrupting the natural form of hand interaction by using additional sensors due to the importance of this natural form of interaction in sensitive aspects such as surgery and medical training.

Physical grasping is an intuitive interaction technique by nature due to the highly complex

structure of the human hand. This complexity offers a unique interplay between the fingers that is not necessarily present in gesture-based interaction techniques. This unique interplay between the fingers results in 33 different grasp types that are classified according to real-world scenarios, and translating this versatile interaction technique to AR could provide interaction designers with more interaction solutions that can potentially aid in developing more intuitive and realistic AR applications across different domains such as art, design, and exposure psychotherapy (Ha et al., 2014). Implementing grasping in AR is also less demanding in terms of user training, this is due to the daily use of grasping in real life that makes users readily familiar with grasping as an interaction technique. This is unlike other interaction techniques where users are required to familiarise themselves with certain gestures that are usually designed to comply with the technology used. For example, users need to train on certain mid-air gestures that are not usually used to interact with objects in real environments in order to be confident in using current AR devices. The interplay between fingers that grasping offers also adds a sense of depth for users during the interaction. This is due to the use of more than one finger that usually wraps around a certain object, this is particularly important in AR as occlusion is widely used as a feedback method, and grasping essentially provides a depth cue to users regarding the position of their grasp in relation to the virtual object using this inherited interplay between fingers. This additional information can be implemented alongside current strong depth cues in AR such as occlusion and would provide additional information for users thus giving users more confidence and control during interaction.

This thesis presents a first analysis of right handed freehand grasping (i.e. physical grasping of virtual objects without the user of any wearable sensors) of two abstract virtual objects (cube and sphere) in exocentric AR using one grasp type that is the medium wrap grasp. Evaluating physical grasping and realising its potential in AR in quantifiable measures will aid in understanding the accuracy, transferability and usability of this interaction technique in AR environments. By highlighting key user errors and behaviours, grasping can be potentially implemented in more intuitive AR applications to bridge the gap between reality and virtuality once its limitations, problems and advantages are realised.

1.2 Research Objectives

The aim of this work is to evaluate and improve dexterous freehand grasping interaction with virtual objects in exocentric AR, that is an AR environment where users can directly interact with virtual objects that are viewed indirectly by users (i.e. through a monitor) in an exocentric manner (i.e users looking at the environment from the outside). This is achieved through the following objectives:

- Define current methodologies used for implementing grasping in AR environments
- Review and define the parameters of physical grasping that need to be considered for

translating grasping into AR

- Report on key problems and human behaviour insights for grasping virtual objects in exocentric AR through measuring the accuracy of this interaction technique
- Measure the impact of additional visual feedback (i.e. dual view visual feedback and drop shadows) on the accuracy and usability of grasping virtual objects in AR
- Quantify the impact of user-based grasping performance measures on grasp performance and usability when used as interaction tolerances during grasping in AR
- Evaluate the suitability of translating physical grasping parameters and rules (e.g. task constraints and grasp phases) to AR

1.3 Thesis Structure

The aim of this work is to evaluate and improve natural dexterous freehand grasping interaction with virtual objects in exocentric AR. Firstly, a review of the parameters and methods of grasping as an interaction technique in both real and virtual environments is presented. Four user studies are then conducted to quantify the accuracy and usability of grasping in exocentric AR (see Figure 1.1 [page 4]).

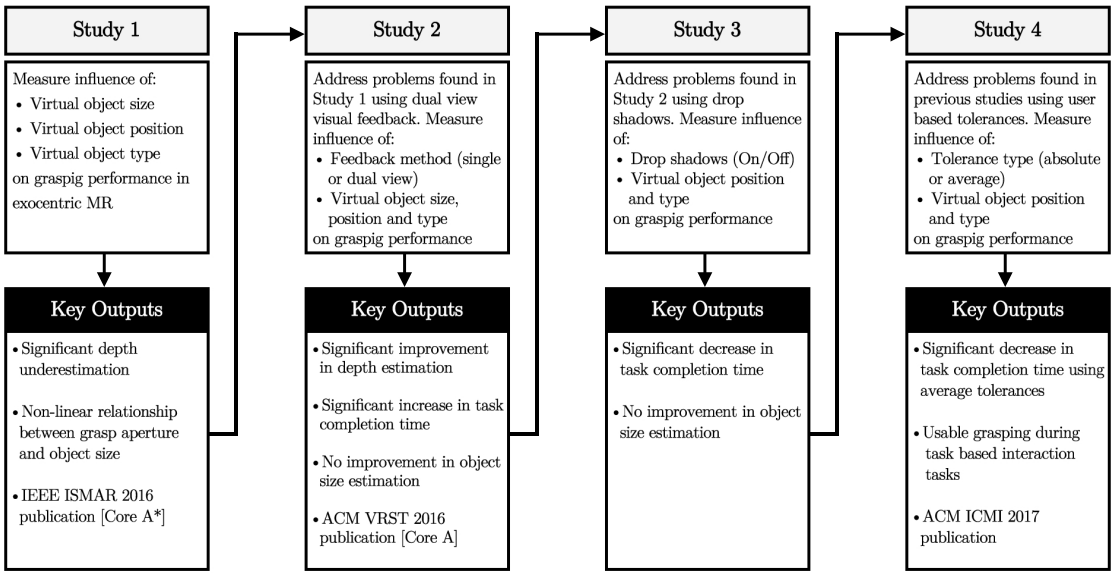


Figure 1.1: Framework of the four user studies presented in this thesis

Findings from user studies are employed to further develop insights to improve user grasp performance in AR. This is presented in eight chapters as follows:
In **Chapter 2**, the background research into interaction in virtual environments is presented. It begins with an introduction of the principal definitions of virtuality and perception. This is followed by an introduction of the definitions of interaction, and a discussion of the two components in the interaction framework with virtual objects in AR environments: user interfaces and

interaction techniques. Finally, current methods for implementing and analysing hand-based interaction techniques are discussed.

In **Chapter 3**, the background research into grasping as an interaction technique is presented. It begins with a detailed discussion of the principal definitions of grasping physical objects in real environments. This is followed by a discussion of the key biomechanical and neurophysiological parameters that enable performing "reach to grasp" movements for real physical objects. A review of the current grasp parameters used to analyse physical grasping such as grasp constraints and phases are then presented. Finally, the concept of freehand grasping of virtual objects in exocentric AR is presented and defined, along with a review of the applications and current problems of this particular interaction technique.

In **Chapter 4** the methodologies and metrics used in this work to analyse grasping accuracy and usability in exocentric AR are presented. It begins with a review of current methods used for analysing human performance in AR environments. This is followed by an outline of the four studies presented in this work detailing the aims, design and variables of each study. Finally, the grasp parameters and experiment design considerations in this work are presented and discussed.

In **Chapter 5**, a first study looking into user grasp accuracy and usability in exocentric AR is discussed. The data collected from two separate experiments in this study is used to quantify the influence of virtual object size, type and position against grasp accuracy using the grasp metrics proposed in Chapter 4. Key problems found in grasping interactions in exocentric AR are then discussed.

In **Chapter 6**, a novel dual view visual feedback method is proposed to mitigate the problems found in Chapter 5. The user study presented in Chapter 5 is replicated with the addition of a second camera to provide additional visual feedback to users. The accuracy of grasping, feedback method and usability using this proposed feedback is then discussed, and user performance problems found using this feedback method are presented. Grasping interaction problems found in Chapter 5 are also re-evaluated through quantifying the impact of the proposed feedback method on grasping performance.

In **Chapter 7** drop shadows are evaluated as an additional visual feedback method alongside occlusion to address problems discussed in Chapter 6. The user study presented in Chapter 5 is replicated with the addition of drop shadows to provide additional visual feedback to users. Grasp accuracy and usability of drop shadows in exocentric AR and problems found using this method are then discussed. Grasping interaction problems found in Chapter 6 are also re-evaluated.

In **Chapter 8**, user grasp performance results from Chapter 5 are used as tolerances during the interaction to improve grasp performance and address the problems discussed in Chapter 7. The user study presented in Chapter 5 is replicated with the addition of user-based grasp toler-

ances triggering visual feedback of grasp attainment. This study aims to investigate how grasping can be usable in task-based interactions without requiring users to be highly accurate in virtual object size and position estimation during grasping movements. The accuracy of grasping and usability of the proposed method is then discussed, and user performance problems found using tolerances are presented. Grasping interaction problems found in previous studies are also re-evaluated.

In **Chapter 9**, a discussion of the findings and insights found in all the four user studies in Chapters 5-8 is presented. Findings in this work are discussed in terms of the two metrics used in this work to assess natural freehand grasping in AR, namely Grasp Aperture and Grasp Displacement. This discussion will present the contributions of this work along with the potential benefits of these contributions for the research community, and revisits physical grasping parameters and discusses its suitability for AR environments.

This work is concluded in **Chapter 10** with a summary of findings from Chapters 5 to 8, followed by a critique of the methods used and recommendations for future work in this area will also be presented.

1.4 Contributions

The primary contribution of this thesis is in the first systematic evaluation of natural grasping for exocentric AR. In achieving this a number of other contributions are made:

- Novel grasp metrics to quantify grasp accuracy and usability in exocentric AR (Chapter 4)
- First study to quantify and evaluate user grasp performance in an exocentric AR system (Chapter 5 and Al-Kalbani et al. (2016*a*))
- First study to improve grasp accuracy and depth perception in exocentric AR using a novel dual view visual feedback method (Chapter 6 and Al-Kalbani et al. (2016*b*))
- Implementation and evaluation of drop shadows as an additional feedback cue alongside occlusion to improve grasp performance and depth perception (Chapter 7)
- Implementation of a novel methodology to assess grasp performance through application of user-based performance tolerances during interaction (Chapter 8 and Al-Kalbani et al. (2017))

1.5 Published Papers

The following papers have been published as part of this work:

- Al Kalbani, M. and Williams, I., 2015, September. Accuracy assessment of free hand grasping interaction in mixed reality. *In Proceedings of the Eurographics Workshop on Visual*

Computing for Biology and Medicine (pp. 205-205). Eurographics Association.**[Core B Ranking]**

- Al-Kalbani, M., Williams, I. and Frutos-Pascual, M., 2016, September. Analysis of Medium Wrap Freehand Virtual Object Grasping in Exocentric Mixed Reality. *In 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 84-93). IEEE. **[Core A* Ranking]**

- Al-Kalbani, M., Williams, I. and Frutos-Pascual, M., 2016, November. Improving freehand placement for grasping virtual objects via dual view visual feedback in mixed reality. *In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (pp. 279-282). ACM. **[Core A Ranking]**

- Al-Kalbani, M., Frutos-Pascual, M. and Williams, I., 2017, November. Freehand grasping in mixed reality: analysing variation during transition phase of interaction. *In Proceedings of the 19th ACM International Conference on Multimodal Interaction* (pp. 110-114). ACM. **[Core B Ranking]**

- Blaga, A.D., Frutos-Pascual, M., Al-Kalbani, M. and Williams, I., 2017, October. [POSTER] Usability Analysis of an Off-the-Shelf Hand Posture Estimation Sensor for Freehand Physical Interaction in Egocentric Mixed Reality. *In 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 31-34). IEEE. **[Core A* Ranking]**

- Al-Kalbani, M., Frutos-Pascual, M. and Williams, I., 2018. Usability Evaluation of Freehand Grasping of Virtual Objects in Exocentric Mixed Reality. *In ACM Transactions on Computer - Human Interaction*. ACM. **[Core A* Ranking - In Submission]**

- Blaga, A.D., Frutos-Pascual, M., Al-Kalbani, M. and Williams, I., 2018. A Grasping Taxonomy for Virtual Objects in Virtual and Augmented reality Environments. *In the International Journal of Human-Computer Studies*. Elsevier. **[Core A Ranking - In Submission]**

Chapter 2

Hand Based Interaction in Mixed and Augmented Reality

2.1 Introduction

This chapter will review and define the different components of virtuality and virtual object hand-based interaction. Interaction is a key component in forming user experiences in Virtual Environments (VEs) and is broadly defined in the context of HCI as the exchange of information between a human and a computer during a task (Hix and Hartson, 1993). In Mixed and Augmented Reality (MR and AR) systems, where a real user is co-located with virtual objects, hand based manual interaction can be defined as the interface between the user's hand and virtual object/s that is mediated by user interfaces that represent the environment and interaction techniques that represent the user. Current research presents various methods of implementing hand based interactions, and these can be classified according to the type of sensory feedback provided to the hand. This classification divides methods presented in the literature into wearable based hand interactions, where sensors are placed on the hand or/and the arm mainly to provide tactile and haptic feedback; and freehand based interactions where no sensors are placed on the hand or arm that more closely resemble hand interactions in real environments.

This work will analyse the freehand form of grasping that is the manual (using one hand) grip between a (real) user and a (virtual) object without the utilisation of wearable sensors. In many applications, this form of interaction is preferable due to the discomfort of wearable devices (Suzuki et al., 2014) and the often time-consuming configuration and user adaptation (Holz et al., 2008) of them. Moreover, other studies (Ponto et al., 2012) have also illustrated that wearable methods of user feedback, notably biofeedback or electromyograms (EMG), can aid in human grasping, but often cause fatigue and discomfort. In addition, analysing this form of grasping in a realistic manual natural interface would aim to recreate the direct interface between a user's hand and a virtual object in a way that replicates a real environment. This will enable users to grasp

virtual objects in the same way that they would grasp real objects in a real environment, and accordingly, allow analysis of this natural form of interaction.

Section 2.2 firstly discusses the different iterations of presenting or displaying virtual environments, where the type of environment developed for this work is outlined and defined based on the virtuality continuum presented in literature (Milgram and Kishino, 1994). Section 2.3 [page 11] defines virtual object interaction and discusses current research concerning user interfaces and interaction techniques. Section 2.4 [page 14] then defines physical grasping in the context of natural interaction techniques along with a review of current sensor based and freehand based grasping methods and applications. Finally, Section 2.5 [page 22] presents the proposed concept of freehand grasping of virtual objects that is analysed in this work, and reviews its rationale, applications and current problems.

2.2 Virtuality Definitions

Perceiving the surrounding world, be it real or virtual, is an important initial step that occurs before any motor interaction takes place. Perception is defined as the acquisition of information from an environment using different sensory organs (such as eyes, ears and fingers) that is then transformed into experiences of sounds, events, objects and tastes (Roth and Frisby, 1986). Perception is a complex process, that involves other cognitive aspects such as memory, attention, language and personal experiences (Rogers et al., 2011).

Early work of Milgram and Kishino (1994) described the terms “real” and “virtual” using the virtuality continuum shown in Figure 2.1 [page 9]. They argued that the real and virtual environments are opposing poles in a reality virtuality spectrum, and should not be considered as alternatives to each other even though they exist as separate entities (Milgram and Colquhoun, 1999).

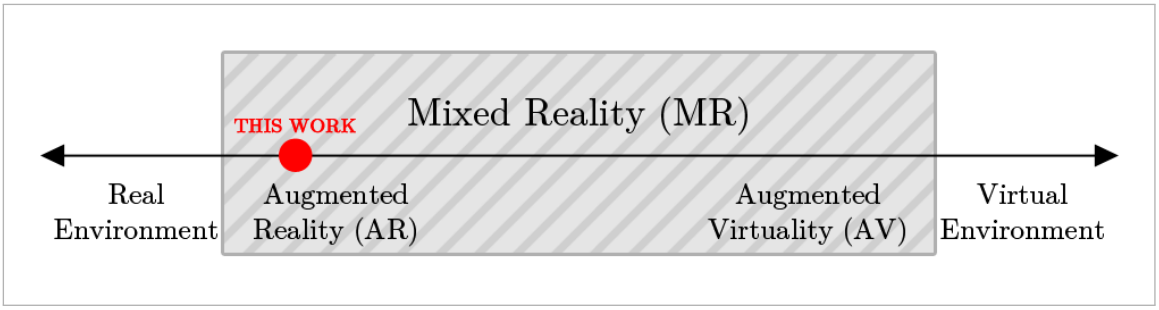


Figure 2.1: Virtuality continuum presented by Milgram and Colquhoun (1999). The red circle indicates the position of this work within this continuum. Image adapted from Milgram and Colquhoun (1999)

Taking into account factors such as scene reality, real video or virtual (computer generated), world view (direct view through air/glass or indirect through a medium such as a monitor) and navigation type or user view of the environment (exocentric or egocentric), this spectrum

clearly classifies the different iterations of presenting or displaying virtual environments. Based on this virtuality spectrum (see Figure 2.1 [page 9]), the following definitions can be deduced:

- **Virtual Object:** a modelled object that is present in essence or effect, but not physically (Milgram and Kishino, 1994), where simulation of the object is required in order for it to be viewed, as in reality; it does not exist
- **Real Environment:** the left extremity of the spectrum that represents the real physical world we live in without any added computer graphics to the environment
- **Augmented Reality (AR):** augmenting a real environment (or elements of it) by virtual (computer generated graphics) objects that coexist in the same space as the environment (Van Krevelen and Poelman, 2010)
- **Augmented Virtuality (AV):** augmenting a virtual environment (or elements of it) by real objects that coexist in the same space as the environment
- **Virtual Environment (VE):** the right extremity of the spectrum that represents a completely synthetic environment built using virtual computer graphics, where the participant feels varying levels of immersion and presence. In this context, Virtual Reality (VR) can be described as the experience of interacting and experiencing a fully artificial environment that makes it feel virtually real (Gigante, 1993), whereas VE describes the general three-dimensional modelling of the environment
- **Mixed Reality (MR):** a subset of Virtual Reality (VR) (Milgram and Kishino, 1994), where real and virtual elements or objects are merged and presented together within a single environment or display, that is anywhere between the two extremities (real and virtual) in the virtual continuum (Milgram et al., 1995). Thus AR and AV are within the bounds of MR.

In other terms, perceiving the type of environment one is in is largely dependent on the amount of virtual object augmentation present in an environment (no augmentation is real, low augmentation where the majority of the environment is real is AR, high augmentation where the majority of the environment is virtual is AV and full augmentation is VR). Realising the different methods of virtuality is important, as using the general VR label for environments that do not necessarily encompass complete immersion, presence and augmentation is common and inaccurate (Milgram and Kishino, 1994). Unlike VR where the user is isolated from the real world, exploring different means of VR (i.e. environments under the MR label) allows users to view the real world, themselves and the virtual objects simultaneously, thus offering a high bandwidth of communication between the user and intuitive manipulation of virtual objects (Billinghurst and Kato, 1999).

The general term MR can be used to describe this work, as the findings in this thesis can be used in different environments that correlate to different classifications in the virtuality continuum such as virtual TV studios (Méndez et al., 2016; Hough et al., 2015) (Augmented Virtuality), or implemented using different motion capture or feedback methods such as HMDs

(Augmented Reality). However a more accurate description of this work conforms to the definition of Augmented Reality (AR) in the virtuality continuum (red circle in Figure 2.1 [page 9]), as the environment in this work is real and is augmented by virtual objects, thus the majority of our environment is real with minor augmentation by virtual objects, thus the term AR is used to describe this work throughout the thesis.

2.3 Virtual Object Interaction

2.3.1 Interaction Definition

With the current large advances in displays, tracking and computational power, AR technologies are now capable of allowing users to interact with and manipulate virtual content in real-time, proving to be useful in a wide range of applications such as education, architecture, medicine and collaboration. Recent advances in technology have also integrated AR in our daily lives, presenting wearable/mobile interfaces where interactions with augmentations on the real world can occur instantly (Maisto et al., 2017).
Generally, interaction and interact are defined in the Oxford Dictionary as:

> **INTERACTION** ***[MASS NOUN]:***
> ***1.0*** Reciprocal action or influence. ***1.1*** Communication or direct involvement with someone or something. ***1.2*** [Physics] A particular way in which matter, fields, and atomic and subatomic particles affect one another, e.g. through gravitation or electromagnetism.
>
> **INTERACT** ***[VERB - NO OBJECT]:***
> ***1.0*** Act in such a way as to have an effect on each other. ***1.1*** Communicate or be involved directly.

Within the context of HCI, interaction is defined by Hix and Hartson (1993) as the exchange of information between a human and a computer during a task for the purpose of controlling or manipulating the computer (interaction by the user) or informing the user (feedback by the computer), where the interaction aims to increase human productivity, ability or satisfaction.

2.3.2 Interaction Framework

Within the virtual realm, interaction with virtual objects is mediated using a User Interface (UI) that represents the environment and an (or multiple) Interaction Technique/s (IT) that represent the user, where the interplay between the two aspects forms the user experience in VEs. These two terms are defined by Bowman (1999) as:

- **User Interface (UI):** the software and hardware that facilitate the interaction between a human and the computer. The UI includes input and output devices, as well as the software architecture of the interactive system

- **Interaction Technique (IT):** a method (such as grasping) by which a user performs a certain task on a computer (a virtual object in this case) via the UI. The method used is usually influenced by the task requirements or instructions, or by the input devices used in the virtual environment

Over the past decade, interaction designers and researchers have merged real and virtual worlds using novel techniques, thus resulting in novel UIs such as mixed reality, augmented reality, tangible and wearable interfaces (Rogers et al., 2011). Innovative ways of interacting with virtual information have also been developed that range from basic gesture and touch-based interaction techniques to the more complex brain-computer interaction. The next two sections will define the different UIs used in AR in current research.

2.3.3 User Interfaces

Early work of Azuma et al. (2001) highlighted the importance of understanding how to present virtual information, and how users would interact with this information in order to build intuitive and effective AR user interfaces (Zhou et al., 2008). Van Krevelen and Poelman (2010) and Carmigniani et al. (2011) classified user interfaces that involve interaction with virtual objects as:

- **Tangible and 3D pointing:** interfaces where users can interact with and manipulate virtual objects in three-dimensional space using real physical objects, such as controllers or paddles

- **Haptic and gesture recognition:** interfaces where users are equipped with sensory devices that provide haptic, tactile or force feedback that aims to aid in gesture recognition to provide more accurate interactions with virtual objects

- **Visual and gesture recognition:** interfaces where hand movements and gestures are tracked visually, without the use of any hand worn devices, using a depth camera or a motion capture device (e.g. Kinect and Leap Motion sensors)

- **Hybrid:** interfaces that combine different, but complementary interfaces that allow interactions through multiple modalities, thus, for example, a user would be able to interact with a virtual object using speech, gaze, hand or a tangible device such as a joystick

- **Collaborative:** interfaces that utilise multiple displays to facilitate co-located and remote hand interactions with virtual objects

- **Brain-Computer Interfaces (BCIs):** interfaces that interpret complex processes of the human brain and user intents using signal processing (Kerous and Liarokapis, 2016), and are defined by Kerous and Liarokapis (2017) as artificial interfaces between the user's brain and a computer that do not require physical movement from the user to control a computer or virtual objects

2.3.4 Interaction Techniques

Interaction with virtual objects can be achieved using a variety of techniques in AR. Interaction designers have incorporated various types of input modalities to mediate interaction with virtual objects for different applications, depending on the interaction tasks required. Billinghurst et al. (2015) classified common interaction techniques in AR as:

- **Information browsers:** augmentation of AR information on the real world, where users browse through information or manipulate the window on which the information is displayed. This technique is mainly used in navigation and simulation AR applications, and is often limited due to the lack of direct interaction with virtual objects or information

- **3D interaction:** direct interaction with virtual objects in space using 3D interaction techniques, such as joysticks and haptic device (e.g. Phantom) that aim to create an illusion of the physical presence of virtual objects. This technique offers various interactivity solutions in AR applications such as training, entertainment and educational AR systems. However a prominent problem in this particular interaction technique is that the methods used are different from those used to interact with physical objects. Meaning that in 3D UIs, users are required to hold or use devices to manipulate virtual objects, whereas as for physical objects in real environments users mainly use their hands for direct manipulation or translation of objects

- **Tangible:** interaction with virtual objects using physical objects that represent virtual information. This technique allows users to intuitively interact with virtual objects through the manipulation of physical objects as they normally do in real environments in a seamless manner. However, tangible UIs often suffer from limitations in display capabilities, and the requirement of physical objects using this technique limits its transferability to wearable and mobile AR applications

- **Natural:** unlike 3D UIs and owing to the vast advances in tracking and display technologies, natural interaction that enables users to directly manipulate virtual objects without wearing any sensors is now more common in AR. Due to these advances, users are now capable of directly interacting with virtual objects using complex techniques such as dexterous hand postures and gestures to a high degree of accuracy. This form of natural interaction is desirable in AR when the use of wearable sensors is not feasible (e.g. medical or live broadcasting applications) nor desirable due to potential fatigue to users

- **Multimodal:** interaction with virtual objects that is mediated through a combination of inputs (e.g. using vision, haptics and sound in one application. This technique aims to enrich interactivity in AR systems, however this form of interaction can again be problematic in terms of requiring users to wear haptic or tangible devices

- **Other:** other interaction techniques explore other body parts and capabilities to mediate interaction with virtual information. For example Hatscher et al. (2017) investigated how gaze and leg interactions can aid physicians interact with medical data when their hands are occupied. Lindeman et al. (2012) also investigated the impact of whistling sound recognition in an AR game. More complex examples of such interaction techniques involve Brain-Computer Interaction (BCI) that was demonstrated by Kerous and Liarokapis (2017) where they presented a novel BCI AR system that enables remote messaging communication using only thoughts using Electroencephalography (EEG) electrodes that were placed on the scalp. The use of Electromyography (EMG) that record and detect electrical activities of muscles was also explored in terms of interaction with virtual objects by Ponto et al. (2012) and Yang et al. (2017).

2.4 Natural Interaction

Grasping can be classified as a hand based natural interaction technique that allows users to directly manipulate virtual objects. The action of grasping is defined as the application of functionally effective forces by the hand to an object for a task, given numerous constraints for the purpose of manipulating, transporting or feeling an object MacKenzie and Iberall (1994). Owing to the different UIs in current AR research, grasping is not always presented in its natural form that is used in the real world. For example grasping can be presented using wearable devices on the hand to recreate the haptic feedback experienced in grasping in real environments, or in a tabletop scenario where grasping is influenced by the tabletop setting. For this reason, the next section will review the current method and applications of grasping in different AR UIs, before introducing freehand grasping in Section 3.1 [page 30], that is the type of natural grasping assessed in this thesis.

2.4.1 Grasping Techniques

In this section, the different techniques and methods for implementing grasping interactions with virtual objects in current research are discussed. Given the focus on this work on grasping, different approaches and techniques for using grasping in AR are classified as wearable based and freehand based, and are defined in this thesis as follows:

- **Wearable based grasping:** grasping that utilises wearable sensors or tracking markers placed on the hand or arm (e.g. tactile feedback) thus the term wearable refers to wearable devices placed specifically on the hand and/or arm. By this definition, any additional wearable devices (such as HMDs) are assumed to be additional feedback methods to enhance the experience in AR

- **Freehand based grasping:** freehand based interaction approaches involve work that involves hand based interactions, ranging from tabletop and gesture-based techniques to

grasping techniques, without placing any additional sensory or feedback devices on the hand or arm

This classification clearly distinguishes between sensory based hand interaction and freehand interaction that is essentially the natural method of hand interaction. The next two sections will review current methods and applications for grasping in different AR UIs.

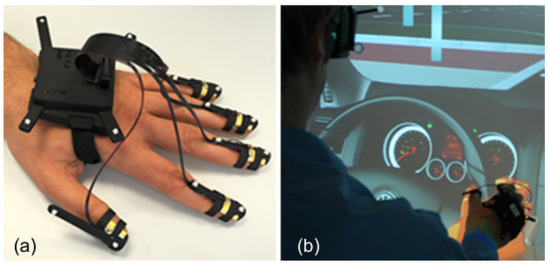
2.4.1.1 Wearable Based

Early work of Bock and Jüngling (1999) measured grasp accuracy using finger trackers (on thumb and index) against grasp aperture changes during single and double step reach to grasp movements, where single step tasks present a virtual disc with a single constant size, and double step tasks change the size of the virtual disc presented instantly or unexpectedly after object presentation. Users were instructed to match their grasp aperture to the sizes of luminous virtual discs displayed 40cm away on a standard PC monitor. Their work has shown that users showed comparable trends in grasp aperture to grasping real physical objects, and also indicated that participants underestimated object size.

Magdalon et al. (2011) assessed the impact of visual and haptic feedback on the kinematics of reach to grasp movements in virtual and real environments. Their work used a 3D tracking system that placed infra-red emitting diodes (IREDs) on the head, trunk, arm, forearm and hand, and additional multiple trackers on the index, thumb, wrist, elbow and shoulder. Three grasp types were assessed in two phases of reach to grasp tasks (reaching and grasping) against spatial and temporal metrics such as grasp aperture, maximum grasp aperture, hand rotation and velocity. A comprehensive analysis in their work reported that hand motion was slower during grasping in VEs with longer deceleration times, this was attributed to the weight of the CyberGlove used. Overestimation of object sizes was also reported for the precision and power grasps assessed, this was attributed to the limited field of view of the HMD used, and the size distortion that is caused by object perception in VEs. Using the same methodology presented by Magdalon et al. (2011), Levin et al. (2015) also assessed reach to grasp and translate movements in physical and virtual environments in post-stroke patients, where they reported that providing haptic feedback in VEs has no impact on reach to grasp movements. Similar to Magdalon et al. (2011), this work also found discrepancies in size estimation (using grasp aperture) of virtual objects due to distorted object distance perception in VEs, and also reported on the decreased velocity of hand movement that was again caused by the CyberGlove used. Tsoupikova et al. (2015) also assessed upper extremity motion functions in post-stroke patients using 10 repetitive reach to grasp tasks or exercises in a VE, that address different aspects of upper extremity motor functions such as reach to grasp, finger individuation and lateral pinch (precision grasp) (see Figure 2.2a [page 16]). Flex and magnetic trackers were used to assess hand/fingers rotation and movements, where a grasp is only considered successful (object sticks to virtual hand if successful) if the hand is in contact with the surface of the object and the joint angles are in alignment with the criteria for an acceptable grasp. Thus for example, if a glass is grasped

too lightly, it slips in an animated fashion in real time, and it animates an explosion if held too tightly. The same process is also implemented for releasing objects. Even though patients found the exercises to be challenging, the authors claimed that patients showed improvement in terms of completion time of tasks after a month of using the system. However, grasp measurement or the grasp criteria chosen were not described.

(a) Assessing upper extremity motion functions in post-stroke patients using wearable based repetitive reach to grasp tasks. Image courtesy of Tsoupikova et al. (2015)



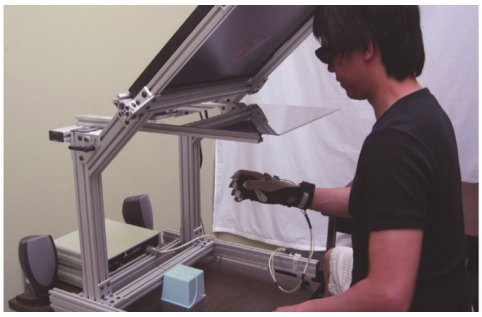
(b) Wearable task based driving simulation to evaluate grasping, manipulating and releasing interactions. Image courtesy of Moehring and Froehlich (2011)

Figure 2.2: Examples of wearable based interactions in AR environments

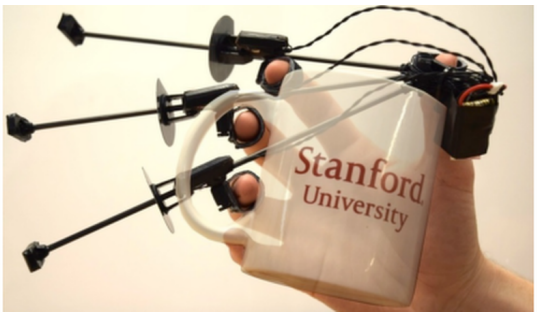
Moehring and Froehlich (2011) used an optical finger grasping tracking system to evaluate grasping, manipulating and releasing interactions in a task-based driving simulation (see Figure 2.2b [page 16]). The tracking system used for the hand was also modified to compare user performance in grasping using different feedback modalities such as visual (using HMDs), pressure based and pinch based tactile feedback modalities. Even though their work showed that performance using a hand-held device was superior in terms of task completion time, they highlighted that using such an indirect interaction is unrealistic and unintuitive. They also suggested that deficits in grasping in VEs in terms of robustness and performance can be compensated using a combination of grasp or pinch detection, that tracks the status of a grasp (object touched, grasped or released) using haptic-based wearable sensors, and a precise finger tracking system. However, no analysis of grasp accuracy was presented.

Virtual grasping, where a virtual hand is modelled and animated according to the real life measurements of a grasp, using a CyberGlove was presented by Borst and Prachyabrued (2013); Borst and Indugula (2006), where they demonstrated the impact of two finger coupling stiffness methods, namely non-uniform and adaptive coupling (see Figure 2.3a [page 17]). Their work presented reach to grasp and translate tasks using 3 grasp types, named by the authors as 2-digit (thumb and index), 3-digit (thumb, index and middle) and 4-digit (thumb, index, middle and ring), where they assessed object position, virtual hand configuration and object motion during the release phase of a grasp. Even though their work is focused on virtual grasping, the approach demonstrated aids in improving the representation accuracy of a grasp virtually.

Choi et al. (2016) developed a mobile wearable haptic device named “The Wolverine” that sim-



(a) Virtual grasping to assess virtual hand configuration and object motion during the release phase of a grasp. Image courtesy of Borst and Prachyabrued (2013)



(b) Mobile wearable haptic device ("The Wolverine") simulates grasping of rigid virtual objects. Image courtesy of Choi et al. (2016)

Figure 2.3: Examples of virtual grasping and mobile wearable interaction approaches in AR environments

ulates grasping of rigid virtual objects, where virtual objects are viewed using a HMD (see Figure 2.3b [page 17]). The mechanical device uses time of flight sensors that track the position of each finger, and an inertial measurement unit (IMU) that provides the orientation of the hand. Their work is currently limited to simulating grasps for objects held in the pad opposition (precision grasps), where the device uses mechanical breaks between the thumb and fingers to simulate grasps of rigid virtual object, meaning that the grasp aperture in a grasp becomes static once the brakes are triggered, thus emulating a physical grasp. This work has shown promising results in terms of tracking the accuracy of changes in grasp aperture and low tracking noise levels, however, it is still unclear whether the mechanical brake mechanism adapted is triggered using collision detection with the surface of the virtual object or is manually triggered.

2.4.1.2 Freehand Based

AR applications present prominent research into freehand interactions. Early research in implementing table-top interactive interfaces (Rekimoto, 2002; Wu and Balakrishnan, 2003) have presented various freehand interaction techniques with objects projected on a surface (see Figure 2.4 [page 18]). Even though those systems presented accurate tracking of hands and measurements of distance of the hand to a surface, only four interaction techniques were studied, one of which employed grasping actions. This was achieved through tracking of multiple fingers that allowed picking up projected virtual objects using the index and thumb fingers. However, no analysis of grasping accuracy was presented, and no attention was given to fundamental grasp metrics such as the grasp aperture.

Early work (Lee and Kim, 2004; Buchmann et al., 2004) also addressed occlusion problems in 2 dimensional AR interactions using freehand interaction techniques such as dragging and dropping (see Figure 2.5 [page 18]). Gestural interactions with virtual objects which included grasp-

Figure 2.4: Examples of early work in tabletop freehand interaction techniques with virtual objects projected on a surface. Images courtesy of Rekimoto (2002); Wu and Balakrishnan (2003)

ing were also developed by tracking the index and thumb fingers. However, early work was limited to studying 6 hand gestures, measuring grasp accuracy was not addressed, and users of those systems highlighted problems of fatigue and tracking inaccuracies. Moreover, both studies indicated that users found both systems intuitive and easy to use without the use of a formal evaluation study.

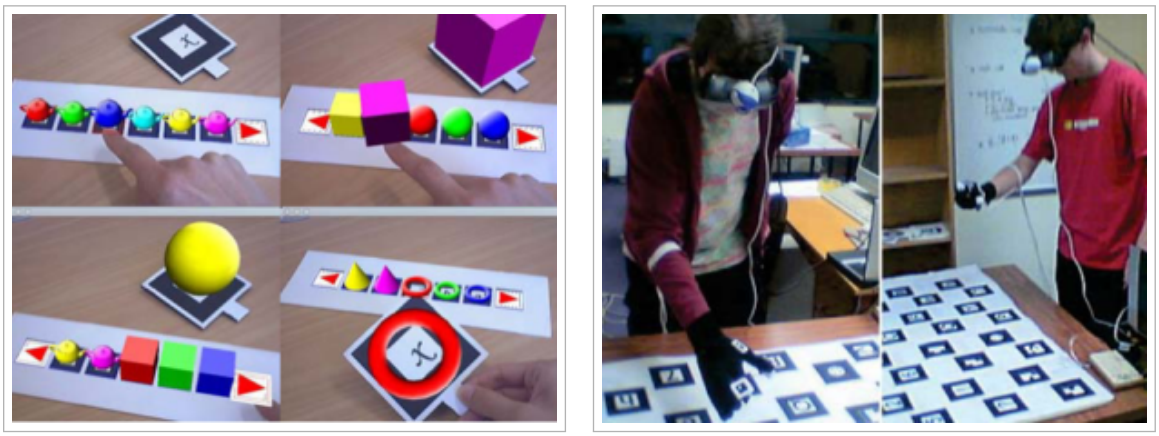
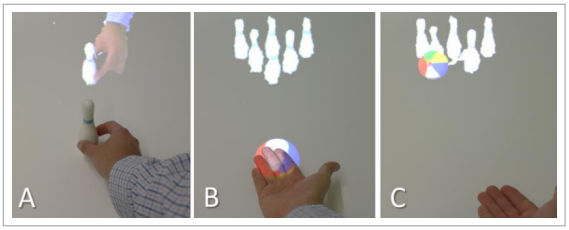


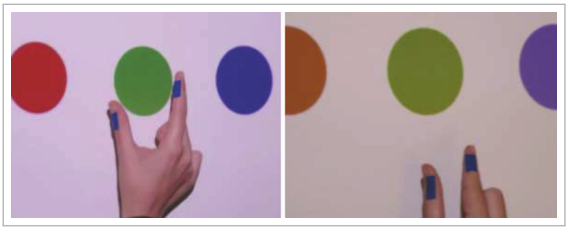
Figure 2.5: Examples of early work addressing occlusion problems and hand gestures (including grasping) in AR. Images courtesy of Lee and Kim (2004); Buchmann et al. (2004)

More recently, freehand interactions in a table-top context were presented by Benko et al. (2012) (see Figure 2.6a [page 19]). Holding, moving and knocking down are the three interaction techniques utilised in this system. Even though a usability evaluation showed users were successful in perceiving projected virtual objects, simulating realistic grasping interactions was not implemented. Hondori et al. (2013) also developed a tabletop AR rehabilitation system, where virtual objects are projected on a table, to assess hand and arm motions in primitive daily tasks such as pointing, grasping, reaching and tilting (see Figure 2.6b [page 19]). This work used computer vision techniques to quantify therapy based performance parameters such as velocity, range, mo-

tion smoothness, grasp aperture and completion time using camera tracking, where coloured markers or stickers (not sensors) placed on the hand were segmented to extract the aforementioned performance measurements. Even though their work reports on performance results of different freehand interactions, including grasping, objective analysis of these interactions was not addressed. The authors also claimed that the developed system was simple to setup and use remotely by patients in their homes, however, the system was presented as a proof of concept by the authors and was only tested on two subjects.



(a) Freehand interactions (holding, moving and knocking down) in a tabletop setting. Image courtesy of Benko et al. (2012)



(b) Tabletop AR rehabilitation system to assess hand and arm motions in primitive daily tasks. Image courtesy of Hondori et al. (2013)

Figure 2.6: Examples of recent tabletop freehand interaction systems

Gestures that emulate human grasping were computed in a 3D handheld AR interface that was developed by Bai et al. (2013) (see Figure 2.7a [page 20]). A depth sensor attached to a tablet allowed acquiring 3D spatial positions of the index and thumb fingers, this information was then used to perform moving, scaling of virtual objects. Even though a formal usability study was employed to assess performance in comparison with 2D touch-based interfaces, time was the only metric used to measure the performance of users. Thus this work was more focused in the usability of the system rather than quantification of interaction types, analysis of interaction accuracy was not addressed. Billinghurst et al. (2014) also developed a handheld AR interface that facilitated freehand gesture-based interactions such as precision grasping to pick and move virtual objects rendered on an image based marker (see Figure 2.7b [page 20]). A SoftKinect depth sensor mounted on a tablet, OpenCV and OpenNI libraries were used to track 3D locations of the fingertips (index and thumb), that served as the interaction points for manipulating virtual objects. The authors reported that a user study showed that users spent more time in finishing tasks using 3D gesture-based interactions than with 2D gesture-based interactions, where users subjectively stressed that there is no significant difference between 3D and 2D gesture-based interactions, in terms of mental stress and naturalness, and that using 3D gesture-based interactions were deemed to be more enjoyable. However, no analysis or results were provided to support these findings. In addition, their work also focused on determining types of gestures users choose or prefer when performing "gesture-in-the-air" based interactions with virtual objects in AR again using image-based markers. This was tested in a tabletop scenario, where a facing down Kinect sensor mounted 100cm above the table facilitated hand tracking, occlusion and 3D reconstruction of the hand in the AR environment that is viewed using a HMD. Users in

this study were shown animations of virtual objects moving and were then asked to determine which gestures could have caused that motion. This showed that users only used a small variety of hand postures, or a variety of the same hand posture and named 11 postures that were frequently chosen by users such as “pinch - fingers together, pinch - fingers spread, grasp - all fingers” as named by the authors, and also highlighted the importance of choosing gestures that reflect real world interactions, and letting users choose the grasp that they desire. Although this work presented valuable recommendations for interaction designers in understanding which hand postures are suitable for hand-based interactions, analysing grasp accuracy was not addressed as the work was disengaged from theoretical work of grasping real physical objects and limited by the tabletop and image-based markers scenario.

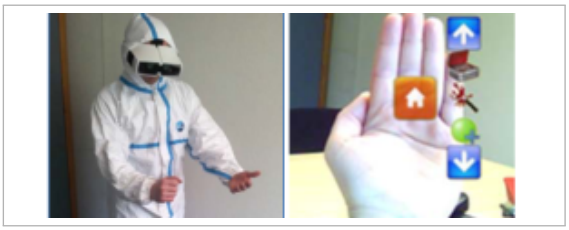
(a) Handheld AR interface to perform moving, scaling of virtual objects. Image courtesy of Bai et al. (2013)

(b) Handheld and tabletop AR interfaces that facilitate precision grasping. Image courtesy of Billinghurst et al. (2014)

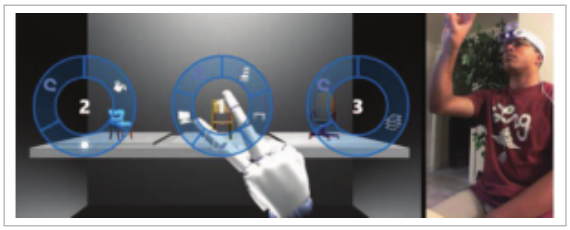
Figure 2.7: Examples of handheld AR that facilitate freehand grasping interactions

Datcu and Lukosch (2013) presented AR freehand interactions in a crime scene investigation application (see Figure 2.8a [page 21]). Even though this work presented novel methods in freehand pointing interactions in an AR context, all of which reflected the dexterity of the human hand, the study only considered 4 natural freehand interactions, and analysis of accuracy was limited to the pointing interaction that is not considered a grasping technique. Davis et al. (2016) presented depth based freehand selection and manipulation of virtual objects with a specific focus on virtual menus (see Figure 2.8b [page 21]). A Leap Motion sensor was used to track hand movements and visual feedback was provided using a standard monitor, and opted against using HMDs due to what they described as technical limitations and instead placed the Leap Motion on the forehead using a headband to make their system transferable in walkable AR environments. This work presented a novel technique in hand based virtual menu selection by adopting a crossing boundary method, thus each menu has two collision zones (one inside the rim of radial menus and another outside), where the two boundaries correspond to two different interactions once a collision is detected with either of them (the first zone highlights the

menu, and the second zone selects an action). However, grasping interactions or accuracy was not addressed in their work, and it is currently unclear which virtual objects were manipulated or selected.



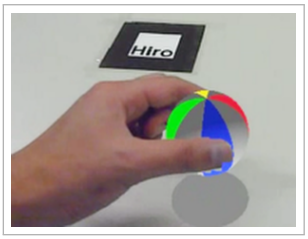
(a) AR freehand interactions in a crime scene investigation application. Image courtesy of Datcu and Lukosch (2013)



(b) Depth based freehand selection and manipulation of virtual menus. Image courtesy of Davis et al. (2016)

Figure 2.8: Examples of freehand interactions and manipulation of virtual menus

Freehand grasping was presented by Suzuki et al. (2014) in an augmented reality context, where wearable devices were excluded from this study due to discomfort (see Figure 2.9a [page 22]). Finger positions in three-dimensional space were detected using a depth camera, those coordinates were then converted to virtual space. Visual feedback was provided using a head mounted display (HMD), to reduce the visual gap between the user and the virtual object. Findings stated that freehand interaction alongside visual feedback increased the feel of grasping. However, no results or in-depth analysis was provided to support this claim. Cidota et al. (2015, 2016) presented freehand interactions, including grasping, in an AR system to assess different upper extremity motor impairments in a serious gaming context (see Figure 2.9b [page 22]). The system comprised of an Optical-See Through (OST) HMD that allowed users to view the augmented game on the real world, a depth sensor that was mounted on the OST-HMD to provide hand and finger tracking and an image based marker to provide alignment between the virtual and real worlds and specify the position of virtual objects in the environment. Requirements of the game developed were outlined following interviews with clinicians and patients, and users were instructed to deliver international mail (virtual boxes) to corresponding destinations (virtual target boxes) while making as few mistakes as possible, for example, if a user is presented with a box that has a picture of the Eiffel Tower on it, then the user is required to grasp the cube using a precision grasp (using thumb and index finger) and then move it to the corresponding target location that is a cube with Paris written on it. This work assessed the usability of the game developed in terms of virtual hand representation, where three different modalities of hand representation were tested (no augmented hand, partially augmented hand and fully augmented hand), and engagement using different game experience parameters such as immersion, annoyance and challenge. Their work showed that users preferred hand augmentations rather than no hand augmentation in terms of usability of virtual hand representations, but reported low engagement scores in terms of game experience due to what the authors described as unnatural grasp movements, as their tracking system required participants to grasp objects with



(a) Freehand precision grasping in AR. Image courtesy of Suzuki et al. (2014)

(b) Freehand grasping in AR to assess upper motor impairments. Image courtesy of Cidota et al. (2015)

(c) Distance free duplication method for manipulating virtual objects in AR. Image courtesy of Jung and Woo (2017)

Figure 2.9: Examples of freehand interaction in AR environments without the use of additional sensors on the hand or arm

the palm facing users (where the sensor is placed). Jung and Woo (2017) developed a distance free duplication method for selecting and manipulating virtual objects in AR (see Figure 2.9c [page 22]). Their work divided interaction into selecting (direct and remote) and manipulation (direct and remote), and used an Oculus Rift HMD to view the AR environment with an attached Leap Motion sensor to track hand movements. This method allowed users to select virtual objects directly if close to them by naturally grasping the object, or remotely if virtual objects are far from the user through a ray-casting technique. Once an object is selected remotely, the target object is duplicated at a close fixed position in front of the user. For manipulation, this work focused on rotating and moving interactions, where direct manipulation is again similar to grasping real object naturally, and remote manipulation allows users to move or rotate the duplicated (after selection) object that shares the same parameters as the original target object, thus if a user rotates the duplicated close object, the target far object will also be impacted accordingly. Even though this work presents a novel technique in freehand interaction, grasp accuracy or interaction accuracy and usability was not addressed, as a usability study for this work was not presented.

2.5 Freehand Grasping of Virtual Objects

Current literature offers evidence for the wide use, benefits and potential problems of hand based virtual object interaction (see Section 2.4.1 [page 14]), however analysing the accuracy of real life grasping in a natural AR environment is still largely unexplored as most of AR research is mainly focused on overcoming software and hardware issues (Dünser et al., 2007).
The work in this thesis focuses on one specific interaction technique that is natural freehand grasping in a visual recognition based user interface, where the accuracy and usability of this interaction technique are analysed using real world and novel accuracy measurements.

2.5.1 Definition

Generally, the term freehand is defined in the Oxford Dictionary as:

> **FREEHAND** ***[ADJECTIVE & ADVERB]:***
> ***1.0*** Drawn or executed by hand without guiding instruments, measurements, or other aids.

In the real world, freehand grasping is defined as the physical manual grip between a human hand and a real object without utilisation of any wearable devices.
Within AR, freehand grasping is the virtual extension to its definition in the real world; it is the manual grip between a (real) user and a (virtual) object without the utilisation of wearable sensors. In many applications, this form of freehand grasping is preferable due to the discomfort of many wearable devices (Suzuki et al., 2014) and the often time-consuming configuration and user adaptation (Holz et al., 2008) of them. Moreover, other studies notably Ponto et al. (2012) have illustrated that wearable methods of user feedback notably, biofeedback or electromyograms (EMG), can aid in human grasping, but often cause fatigue and discomfort in users.

User Interface

Freehand grasping is considered a natural real world interaction technique that is used on a daily basis, thus developing a natural user interface is reasonable to facilitate and analyse this form of natural interaction. For this reason, a natural user interface (NUI) is implemented for this work, that is defined as an interface that enables users to interact with a computer in the same way that they interact with the real physical world through using their voices, hands, bodies and speech (Rogers et al., 2011), to analyse this natural form of interaction.
NUIs can potentially be prone to technical problems such as self occlusion, speed/frame-rate (Corera and Krishnarajah, 2011; Rautaray and Agrawal, 2015) and tracking reliability if using vision based tracking devices (Cidota et al., 2016) or computer vision image based techniques to mediate hand or gesture based detection and interaction (Pham et al., 2018) (see Section 2.5.4 [page 26]). However, developing a NUI that mediates this form of natural interaction (i.e. physical grasping) is necessary for this work to give users more control over virtual information in a way that feels more realistic, and to provide an intuitive method of interaction (Hondori et al., 2013), this is in alignment with the current growing interest in developing AR user interfaces that are usable and accessible to users with a wide range of needs, skills and expectations (Oliveira et al., 2017). Furthermore, a NUI can increase the usability and effectiveness of a system as it enables users to use and view their hands directly in the same space as virtual objects (Klein and De Assis, 2013), this is particularly important in grasping interactions as vision plays a major role in forming a grasp strategy. The NUI in this work excludes the use of any wearable devices, even if not placed exclusively on the hand or/and arm (e.g. HMDs), to avoid any biased results or perceptual problems that may be caused due to the use of such wearable devices. Thus the action

of grasping virtual objects will be analysed in an environment that resembles that of grasping real physical objects, where users directly view their hands and are naturally not mounted with cumbersome wearable devices or restricted by limited interaction space while grasping objects.

Interaction Technique

Freehand grasping is a subset of freehand based interaction methods, and the wide use of grasping as an interaction technique in wearable and freehand based AR systems is evident in the literature (see Section 2.4.1 [page 14]). While the current literature offers valuable analysis of grasping in a wide range of applications, analysing the fundamental problems of freehand grasping in terms of accuracy and usability is still largely not addressed. Grasping in current research is usually implemented as an element of a bigger application, where the analysis is mainly focused on assessing a certain application, with grasping in it, rather than the accuracy of grasping on its own. Furthermore, current research that includes freehand grasping does not take into account grasp types that are classified according to comprehensive taxonomies, and aspects of grasp planning that influence grasp formation.
Measuring grasp accuracy is important owing to the fact that the hand is a dexterous tool that can perform a wide range of different interactions (Arkenbout et al., 2015), however, grasp accuracy measures are still unexplored in literature. Furthermore, grasping in AR also lacks a unified grasp taxonomy, where the majority of grasps currently introduced are specific to limited applications and at times biased by environment structure (e.g. tabletop and handheld AR applications). In this work, real-life grasping is recreated in an AR environment using theoretical grasp types, and freehand grasping is analysed as an interaction technique taking into account accuracy, usability and dexterity of a grasp in AR. This will potentially highlight the different usability and perceptual problems associated with grasping in AR.

2.5.2 Rationale

While grasping is one of the primary forms of manual interaction used by humans, the dexterous versatility of the human grasp poses many challenges within virtual object interaction and as such the objective quantification of these problems is largely unexplored. Early review by Moeslund et al. (2006) indicate that vision-based human motion analysis is a thriving area of research that is driven by its potential applications in a wide range of applications (e.g. entertainment and surveillance). Early work of Buchmann et al. (2004) also recommended that AR interfaces enable freehand interactions with virtual objects, as this eases the transition between interacting with real and virtual objects simultaneously by allowing natural and intuitive interactions.
Wearable devices can potentially offer more accurate tracking, however, such devices limit natural movements of users and can be inconvenient to use. Using the hand in its natural form (also described as “Bare Hand”) maximises manipulation of virtual objects and offers the most natural and intuitive forms of interaction (Jung and Woo, 2017). Vision-based AR systems are

considered more natural to use than glove based interfaces (Billinghurst and Buxton, 2011), and are capable of offering non-intrusive robust tracking of the body and hand (de La Gorce et al., 2008), where using the hand as a natural tool for interaction excludes the need for physical markers and cumbersome wearable devices, and users feel comfortable in using their hand as the main tool for interaction (Lee et al., 2008).

The motivation behind analysing the accuracy of freehand grasping in a natural AR is to bridge the realism and naturalness gap between grasping in real and virtual environments, as currently grasp planning aspects such as hand, task and object constraints and their impact on perception and interaction performance on virtual objects are unexplored. Furthermore, hand based dexterity and capabilities in interaction such as the interplay between the index, thumb and middle fingers and their impact on interaction in AR is currently not addressed. Benko et al. (2012) argued that current interactive AR systems largely suffer from "impoverished" input from the real world, as the majority of current AR solutions are mainly focused on output technologies such as HMDs or handheld displays. Furthermore, they also add that users are overburdened with wearable sensors on body and hand, and are unable in current systems to perform the fine motor skills, such as grasping, that is usually used in the real world. Based on this, this work aims to improve input interactions from the real world (i.e. grasping) in AR environments by evaluating and quantifying the accuracy of grasping in an AR environment, this analysis is based on widely accepted theoretical analysis of physical grasping of real objects. This can potentially aid in a better presentation of grasping in terms of accuracy, type and spatial positioning in AR, and also better realise the full of potential of grasping in AR.

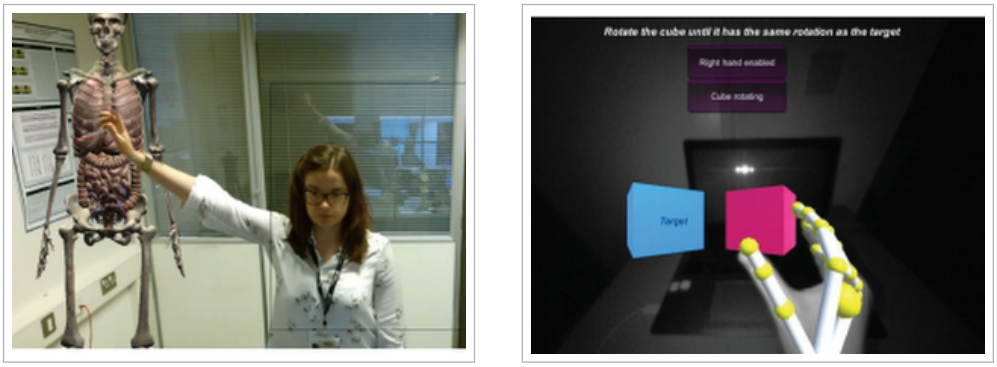
2.5.3 Applications

Grasping is used as an interaction technique in a wide range of AR applications and user interfaces. Freehand grasping, in particular, offers a widely available cost-effective solution to interaction in AR as real-time markerless depth sensors are relatively cheap and offer an acceptable balance between cost and usability (Kitsikidis et al., 2014), this is particularly useful for applications that require sensors to be affordable, portable and easily configured as in medical AR applications (Hondori et al., 2013).

Freehand grasping is also valuable where wearable devices are not valid to use due to the nature of the developed application, for example, the AR crime scene investigation application developed by Datcu and Lukosch (2013) relied on freehand gestures where wearable gloves and markers were not valid to use per the requirements of the application as such devices are not easily exchangeable, and the use of a HMD in their application also excluded the possibility of using additional wearable sensors as that would make the system cumbersome to use by crime scene investigators. Cidota et al. (2015, 2016) also adopted freehand interactions through markerless tracking in their medical application assessing upper motor impairments in stroke patients based on recommendations by members of the clinical community, where the use of wearable sensors was not feasible due to their potential interference with natural body motion,

and freehand interactions (including grasping) in this case offered a more meaningful assessment of upper limb impairments through a natural way of interaction with virtual objects that resembles interaction in real life. Freehand interaction is also required in live TV virtual studios, where presenters are not allowed to use any wearable sensors to aid in interaction with virtual objects, as they can break the illusion of interaction realism with virtual objects for third person viewers. Hough et al. (2015) addressed this problem by implementing bimanual freehand interactions for TV virtual studios, and also offered solutions to increase the fidelity of interaction in TV virtual studios.

The freehand grasping metrics and findings in this thesis build on the previous work in TV virtual studios by Hough et al. (2015), and have also already been applied and validated in a medical application to interact with and visualise different anatomical information (see Figure 2.10a [page 26]). In addition, findings in this work were also used to assess the accuracy and usability of off the shelf sensors (Leap Motion) for grasping in AR environments (Blaga et al., 2017) (see Figure 2.10b [page 26]).



(a) Freehand grasping in a medical application to visualise various anatomical information

(b) Freehand grasping metrics used in assessing usability and accuracy of off the shelf motion capture sensors. Image courtesy of Blaga et al. (2017)

Figure 2.10: Examples of freehand grasping applications in AR environments

2.5.4 Problems

Freehand grasping is potentially subject to some of the various problems in AR environments such as technically limited tracking and perceptual accommodation mismatches. Problems, trends and accuracy of freehand grasping in exocentric AR are unclear as this is currently unexplored in literature. However owing to the fact that freehand grasping is a vision based interaction technique that is facilitated using depth sensors, much research has highlighted key technical problems in the design of hand pose estimation systems. Tracking the dexterous human hand with its more than 20 degrees of freedom can be problematic, and the review of Zhou et al. (2008) highlighted the complexity of the scene and the motion of tracked objects and their degrees of freedom as the main difficulty of real-time tracking of the hand.

Problems in visual hand tracking are summarised by Sturman and Zeltzer (1994) and Erol et al. (2007) as: limited resolution and FOV of camera and depth sensors, insufficient frame rates for rapid hand motion, uncontrolled environments (e.g. varying lighting conditions) and highly articulated structure of the hand and self-occlusion problems of the fingers. Such limitations need to be taken into account when developing and assessing freehand grasping to avoid potential false results and inaccurate interactions.

2.6 Summary

This chapter first defined the different components of virtuality and virtual object interaction, and then reviewed current grasping interaction techniques used in different user interfaces. Understanding current methods of hand based interaction techniques in varying levels of virtuality is important to the understanding of the current problems and limitations of the more complex interaction technique that is grasping. Current research presents various methods and applications for grasping interactions, where grasping is widely used due to its naturalness and effectiveness in representing and measuring human performance. For example, grasping is particularly useful to assess human motor functions in task-based virtual systems.

Current methods that implement grasping in AR environments are divided into two categories: wearable based and freehand based grasping. Wearable based grasping methods use hand based sensors to specifically track one or multiple fingers, and to provide users with the haptic feedback that is naturally experienced during grasping in real environments. Grasp accuracy in wearable based methods is mainly measured using the physical grasping metric that is grasp aperture, defined as the distance between the index finger and thumb. However, they are still problematic due to the inconvenience of wearable sensors that can cause fatigue and influence grasp performance (e.g. users perform grasping slower due to the additional weight of hand based gloves or sensors). In addition, other physical grasp parameters such as grasp type usability and task constraints are not addressed using these methods. Freehand based methods avoid using hand based wearable sensors to enhance usability, and mainly track the hand using visual recognition devices such as HMDs and Microsoft's Kinect. The accuracy of grasping in freehand based approaches is largely unexplored as research using these methods is mainly focused on the usability of the end application developed and not the interaction technique. This is evident by the wide use of task completion time as a performance metric in these systems. In addition, physical grasping parameters such as grasp choice in current freehand based research is mainly influenced by the tracking capabilities of the technology used, thus grasp types used are ones that can be tracked accurately within a certain application and this often leads to unnatural grasp movements.

This chapter also introduced the proposed concept of freehand grasping of virtual objects that extends the definition of physical grasping to AR environments, and is defined as the manual grip between a (real) user and a (virtual) object without the utilisation of wearable sensors. Free-

hand grasping aims to recreate the naturalness of physical grasping in a natural user AR interface where users will directly grasp virtual objects in the same manner as they grasp physical objects. Analysing the accuracy and usability of freehand grasping will potentially aid in the development of AR systems where the use of wearable sensors is not feasible such as medical (Hondori et al., 2013) and live broadcasting (Méndez et al., 2016) applications.
Current grasp methods in AR environments, be it wearable or freehand based, still largely ignore physical grasping parameters. As such the definitions of these parameters and their impact on grasp accuracy is still not clear. This thesis focuses on measuring the accuracy and usability of physical grasping in exocentric AR without the use of any wearable sensors in a natural user interface that resembles grasping in real environments. For this, the next chapter will describe the theoretical background behind the dexterity of the human hand, and discuss physical grasping parameters that are largely ignored in current research, to better understand this grasping interaction technique before assessing it in AR.

Chapter 3

Grasping Interaction

3.1 Introduction

The human hand is a powerful and dexterous tool that mediates the majority of mechanical interactions with our surrounding world (Winters and Crago, 2012). The high dexterity of the human hand makes grasping one of the primary forms of interaction between humans and the surrounding world. A grasp is defined as every static posture where a certain object can be held securely using one hand, however, this definition is representative of only the final stage or goal of a grasp that is holding an object securely. The action of grasping, also known as prehension, extends beyond this definition and is defined as the application of functionally effective forces by the hand to an object for a task, given numerous constraints for the purpose of manipulating, transporting or feeling an object.

The action of grasping is a complex process that begins in the human brain prior to any motor action by the hand. There is currently a large body of research that discusses the action of grasping through analysing not only the structure and biomechanical features of the hand but also the antecedent role of the human brain that allows the human hand to choose the most suitable grasp trajectory and type in relation to an object. This complex process starts in the Central Nervous System (CNS) which then directs the biomechanical capabilities of the hand to reach for and finally grasp an object securely.

Physical grasping is subject to various parameters that play a role in performing a grasp action, most notably grasp constraints and phases. These parameters determine the most feasible grasp type for a given object depending on the task constraints and the stage of the grasp. Current research clearly details these parameters for physical grasping, and also presents various grasp taxonomies that classify possible grasp types depending on task requirements, object shape and hand structure. Understanding these parameters and types and how they can be translated to AR will aid in designing new methods and metrics to measure grasping accuracy in AR, given the lack of metrics and grasp taxonomies to analyse grasp accuracy in current literature for AR environments.

This chapter will define physical grasping, describe the underlying structure of the hand that mediates the action of grasping and discuss the different parameters of physical grasping. Firstly physical grasping and the action of grasping are defined in Section 3.2 [page 30]. The fundemental biomechanical and neurophysiological features and processes that mediate the action of grasping are then described in Section 3.3 [page 30]. Section 3.4 [page 34] reviews widely used grasp taxonomies and methods in classifying grasp types. Grasp planning in terms of grasp constraints and phases is then discussed in Section 3.5 [page 41].

3.2 Grasping Definition

The human hand is a powerful and dexterous tool that mediates the majority of mechanical interactions with our surrounding world (Winters and Crago, 2012). The evolution of the human brain has shaped grasping into a core cognitive ability (León et al., 2014), and one of the primary forms of manual interaction between humans and the physical world that is inherent to human beings (Supuk et al., 2011) as one of the fundamental interactions and essential for performing activities in daily living.
A grasp is defined in the Oxford Dictionary as:

> **VERB** ***[WITH OBJECT]:***
> ***1.0*** Seize and hold firmly. ***1.1*** Get mental hold of; comprehend fully. ***1.2*** (grasp at) [no object] Try to seize hold of. ***1.3*** Act decisively to the advantage of (something)
>
> **NOUN:**
> ***1.0*** A firm hold or grip. ***1.1*** A person's power or capacity to attain something. ***1.2*** A person's understanding

Feix et al. (2009) define a grasp as being every static posture at which an object can be held securely with a single hand. However, this definition can be limited in this body of work as it excludes the grasp stages that occur before (reach) and after (transport) establishing a static secure posture around a specific object. MacKenzie and Iberall (1994) defined prehension, the action of grasping, as the application of functionally effective forces by the hand to an object for a task, given numerous constraints for the purpose of manipulating, transporting or feeling an object. This definition is more representative of where all stages of grasping are discussed and analysed, as it addresses the motor and task aspects of grasping interactions.

3.3 Grasping Physical Objects

3.3.1 Biomechanics of Grasping

The human hand is capable of extremely dexterous interactions and postures owing to the various skeletal and muscular degrees of freedom (Nowak and Hermsdörfer, 2009). The hand consists of five digits (fingers) that are built using a collection of bones, tendons, muscles, ligaments,

fascia, and vascular structures enclosed by skin (MacKenzie and Iberall, 1994). The dense central nervous system utilises the thousands of peripheral nerves in muscles, skin and joints to mediate dexterous postures and interactions.

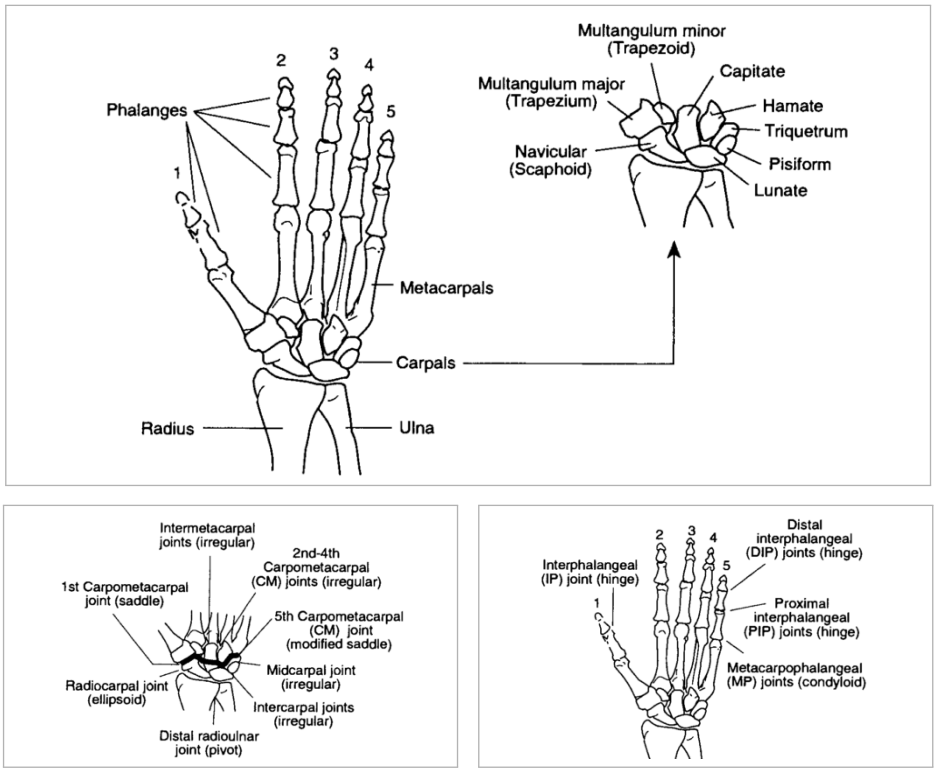


Figure 3.1: Human anatomy of the hand and wrist. The human hand consists of 8 carpal bones in the wrist, 5 metacarpal bones in the palm, 2 phalanges in the thumb and 3 phalanges in each of the four fingers

The human hand consists of 27 bones, 39 muscles (Nowak and Hermsdörfer, 2009) and 28 degrees of freedom (MacKenzie and Iberall, 1994). The hand consists of 8 carpal bones in the wrist, 5 metacarpal bones in the palm, 2 phalanges in the thumb and 3 phalanges in each of the four fingers (León et al., 2014; MacKenzie and Iberall, 1994) as shown in Figure 3.1 [page 31]. Much of the versatility of the hand is due to the support provided by the wrist and arm (Cutkosky, 2012). The wrist also provides the precise orientation of a grasp through joint movements between wrist bones and the forearm (Bennett and Castiello, 1994) and studies have shown that the majority of users prefer using the wrist to control grasping as it is considered a natural way to perform grasping tasks faster (Chapin and Moxon, 2000). This support, even if static, is required to perform a stable grasp.

Due to its dexterity, the human hand is capable of grasping objects of different sizes and shapes, in a manner that can be forceful or delicate depending on the intended task (Bennett and Castiello, 1994). For example, the human hand will exert a low force to pick up a needle using a precision grasp, and a higher force to grasp a pint of water using a power grasp. This flexibility in applying different forces and grasp structures is facilitated by the finger muscles (Cutkosky, 2012). Stud-

ies have shown that at least 5 muscles are required to perform lateral (precision grasps for small objects) and palmar (power grasps for large objects) grasps (Chapin and Moxon, 2000), where the intrinsic muscles within the hand stabilise fingers for fine manipulation, and the extrinsic muscles within the forearm provide most of the force for grasping heavy or large objects.
The size of the human hand and its bones are relatively small (MacKenzie and Iberall, 1994), this mechanical design of the hand gives it a higher bandwidth for mobility, adaptivity and control (Winters and Crago, 2012) that allows the hand to perform both small and large deformations when required. As the fingers and thumb circle around an object, the tissues of finger pads and the palm adapt to the surface of the object being grasped (Cutkosky, 2012).

3.3.2 Neurophysiology of Grasping

The mechanical structure of the hand plays a major role in performing a grasp, where the different hand muscles direct the bones in certain configurations to create a physical grasp posture (MacKenzie and Iberall, 1994). However, the main controller of the mechanical capabilities of the human hand is the central nervous system (CNS) that facilitates interactions between the brain and the rest of the human body. The CNS comprises of the brain and spinal cord, where the spinal cord acts as a two-way carrier of neural messages from the brain to muscles (for motor functions) and skin (for sensory functions) and sends signals about the rest of the body to the brain (Anderson, 1985; Helander, 2014).

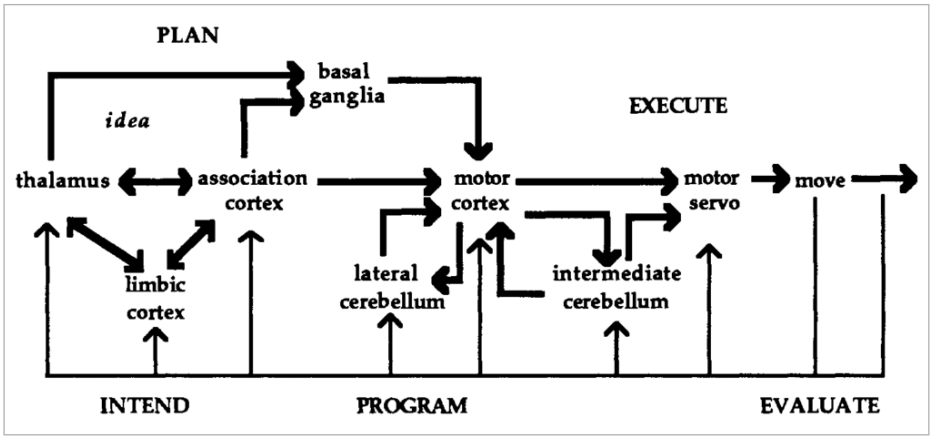


Figure 3.2: CNS areas involved in planning, programming and execution of grasping movements. First: the cerebral cortex, basal ganglia and cerebellum plan the grasping action prior to any movement. Second: the association cortex, limbic cortex, lateral cerebellum and basal ganglia send the intended/planned motor commands to the motor cortex. Third: the motor cortex then executes movement (trajectory planning). Fourth: sensory information is then fed back to cortical areas for further planning (e.g. grasp corrections). Image courtesy of MacKenzie and Iberall (1994)

Understanding the underlying processes by the CNS in performing a grasp is important for this work to better interpret the impact of different aspects of a grasp that are dictated by the CNS

such as grasp strategy choice, grasp phases and grasp motion on user performance. In addition, understanding how a grasp normally functions under the rules of the CNS will also help in better understanding the impact of the missing haptic feedback in this work on user freehand grasping performance, that is an integral from of feedback to the CNS processes in forming a successful grasp for physical objects. The evolution of the human hand has been paralleled by the significant changes in the CNS, where the motor and sensory cortical areas that are devoted to the hand has been largely expanded (Nowak and Hermsdörfer, 2009). Thus even in the simplest grasping tasks, the CNS is capable of using billions of nerves in different anatomical regions to encode complex motor planning and sensory functions from the muscles and skin in the human hand (MacKenzie and Iberall, 1994; Winters and Crago, 2012; Bennett and Castiello, 1994).

The CNS is unique in its capability to control a wide range of tasks that range from walking and jumping to complex fine manipulation skills such as grasping (Winters and Crago, 2012). Successfully grasping an object represents the end result of a motor sequence that starts with complex motor and sensory planning in the CNS way ahead of the grasping action itself (Bennett and Castiello, 1994). Allen and Tsukahara (1974) presented a model of the CNS structures that are involved in planning, programming and execution of grasping actions (see Figure 3.2 [page 32]).

The different phases of a grasp will be explained in more detail in Section 3.5.2 [page 43], but an example of a reach to grasp movement will be given for the purpose of explaining the functionality of the different structures and pathways of the CNS during grasping interaction. For example, if a user is required to grasp a ball that is placed on a table, the first point of contact that the user will have with the object at this point is solely visual. Once the user intends to reach for the object, the brain uses this visual information to plan a specific posture, using the limbic cortex, thalamus, motor cortex and basal ganglia, that is suitable for the object being grasped as shown in Figure 3.2 [page 32] (Allen and Tsukahara, 1974; MacKenzie and Iberall, 1994). Every able bodied human is equipped with motor rules that are progressively installed in the CNS as it matures (Jeannerod, 2006), and these rules are activated by the CNS once an action has to be performed to achieve a certain goal. Thus in this case of picking a ball, the posture and movement chosen by the CNS will be largely dependent on the motor rules of the user. Once the posture is chosen by the CNS, the motor cortex sends a signal through the dorsal column nuclei in the spinal cord (passing by the thalamus) to the efferent pathway which then alters the movement of the muscle (see red line in Figure 3.3 [page 34]) (Kalaska et al., 1983). After contact is made with the object being grasped, the skin (if grasping a real physical object) or the visuo-sensors (if grasping a virtual object) will send a signal through the afferent pathway back to the motor cortex to address any sensory changes, wrong postures or risky postures such as slipping (see blue line in Figure 3.3 [page 34]).

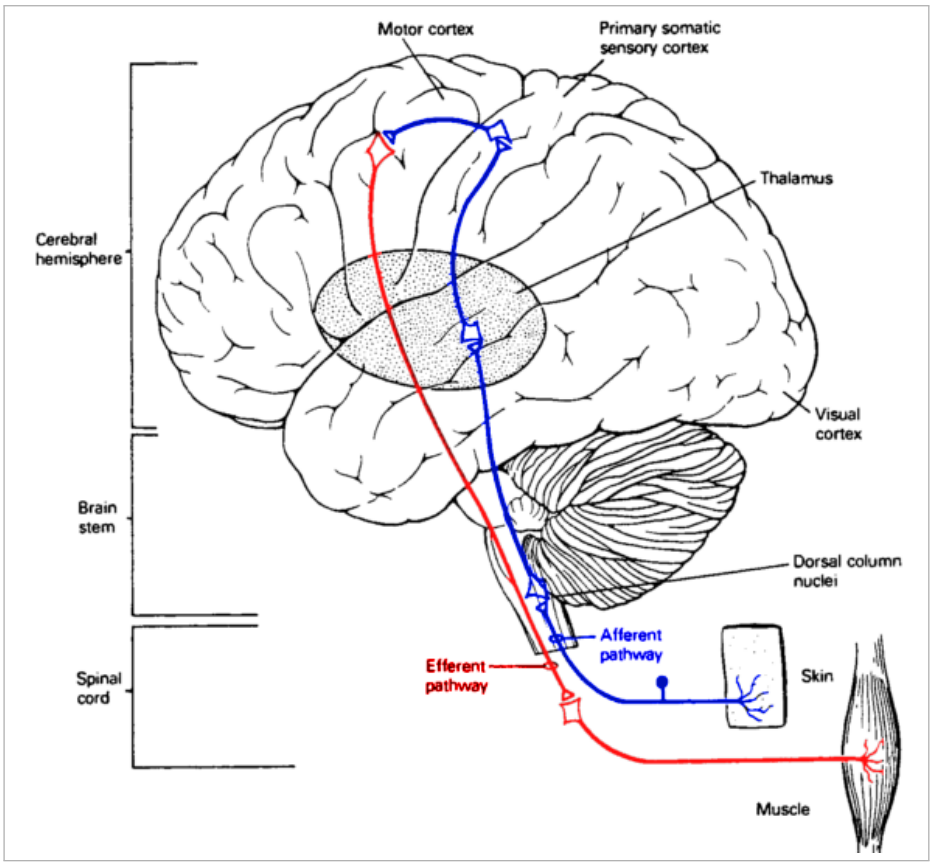


Figure 3.3: Areas of the CNS that are involved in grasping movements. Image courtesy of Kelly et al. (1985)

3.4 Grasp Classification

3.4.1 Descriptive Taxonomies

Bowman (1999) defines the word taxonomy as the "science of classification" and a "specific classification". Derived from the complexity and physiology of the human hand, the grasping process requires various simplifications through the formation of taxonomies to make it easier to understand (Cutkosky and Howe, 1990). Grasp taxonomies are introduced in various domains such as anthropology, hand surgery, hand rehabilitation, robotics, developmental psychology and virtual environments (MacKenzie and Iberall, 1994), and aim to offer a better understanding of the human grasping capabilities (Feix et al., 2009). The high number of degrees of freedom in the human presents a problem in fully understanding grasping capabilities (Nowak and Hermsdörfer, 2009), and grasp taxonomies transform this complex problem into simpler problems by taking into account key factors in choosing a grasp, such as object properties (shape and size), task and opposition type (part/parts of the hand that are applied on the surface of an object during a grasp).

Schlesinger (1919) introduced a first simple taxonomy in 1919 to classify grasping actions and functionality for prosthetic arms, that were used for upper limb injuries in World War I (see

Figure 3.4 [page 35]). This classification approach focused on treating the hand as a tool and determining what type of functionality is required by the human hand to grasp objects of different shapes. Taking into account object shape (cylindrical or spherical), hand shape (hook, open fist, close fist), and hand surfaces (tip, palmar, lateral), six grasp types were introduced in this taxonomy: cylindrical (for cylindrical objects such as a beer mug), tip (for small objects such as a needle), hook (for heavy objects such as suitcases), palmar (for flat thick objects such as a match box), lateral (for flat thin objects such as a piece of paper) and spherical (for spherical objects such as a ball).

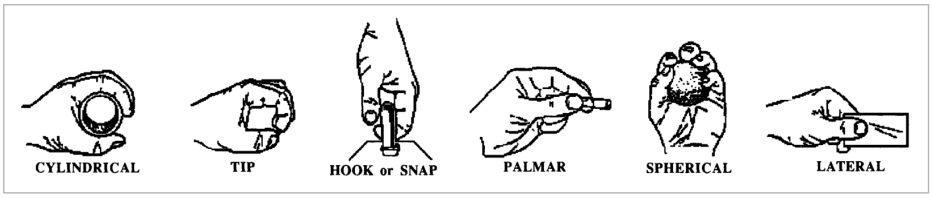
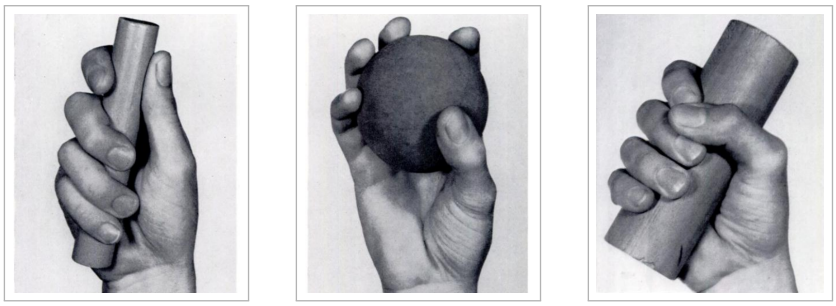


Figure 3.4: First grasp taxonomy by Schlesinger (1919) that treated the hand as a tool, and took into account object shape (cylindrical or spherical), hand shape (hook, open fist, close fist), and hand surfaces (tip, palmar, lateral). Image courtesy of Schwarz and Taylor (1955)

An alternative approach to classifying grasp types was presented by Slocum and Pratt (1946) to better understand the loss of functional hand use due to injuries, where they focused on the opposition parts of the hand to the fingers. This approach reduced the various postures presented by Schlesinger (1919) to three functional components of the hand, and presented three grasp types: grasp (coupled action between the fingers and the opposite palm and thumb of the hand), pinch (thumb pad against pads of the opposing fingers) and hook (flexed fingers where their pads are parallel and marginally away from the palm) (Slocum and Pratt, 1946; MacKenzie and Iberall, 1994).

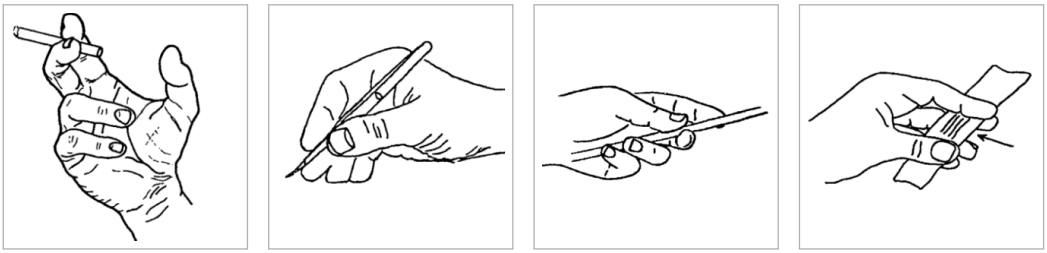
Even though the taxonomies formed by Slocum and Pratt (1946); Schlesinger (1919) were insightful and extensive, task requirements, that are considered to be an important influence on grasp choice, were not taken into account. Napier (1956) argued that the taxonomy of Slocum and Pratt (1946) was not clear and extensive, and thus his work focused on forming a taxonomy that takes power requirements (power or precision) into account, and is based on both the functional and anatomical features of grasping (see Figure 3.5 [page 36]) (MacKenzie and Iberall, 1994). His work presented a detailed anatomical description of the power grasp (see Figure 3.5a [page 36]), where the thumb is adducted, and the fingers flex in opposition to the palm with degrees of freedom that are dependent on the dimensions of the object. He also noted that a level of precision in a power grasp is dependent on the placement of the thumb, where some precision can be achieved if the thumb is adducted, and no precision (maximum power) if the thumb is abducted, this turns into the “coal hammer” grasp type in his taxonomy (see Figure 3.5c [page 36]). Napier (1956) also presented a detailed description of the precision grasp, where the thumb is abducted, and fingers are flexed and abducted (see Figure 3.5b [page 36]).



(a) Power Grasp (b) Precision Grasp (c) Coal Hammer Grasp

Figure 3.5: Napier (1956) presented a taxonomy that focused on task requirements (power or precision), and argued that the hand could form either a power or precision grasp to match any task requirements. Images courtesy of Napier (1956)

Landsmeer (1962) revised the precision grasp definition by Napier (1956), and introduced precision handling, that takes into account the dynamic aspects of precision grasp movements in fine translations and manipulations. A power grasp reaches a definite static phase once an object has been grasped, whereas a precision grasp does not, thus Landsmeer (1962) suggested that grasping an object between the thumb and finger pads facilitates a higher variety of movements as the fingers are capable of mediating such movements.



(a) The adduction grasp. Image courtesy of Kamakura et al. (1980)

(b) External precision grasp. Image courtesy of Patkin (1981)

(c) Internal precision grasp. Image courtesy of Patkin (1981)

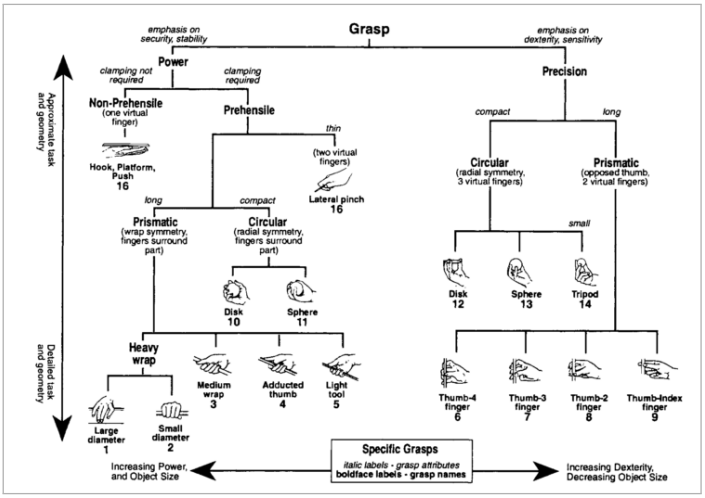
(d) Double grip. Image courtesy of Patkin (1981)

Figure 3.6: Examples of dynamic grasps: Landsmeer (1962) revised the definition of a precision and highlighted the importance of the interplay between the index, thumb and middle finger in the human hand. This lead to presenting new grasp types, that are variations of this interplay

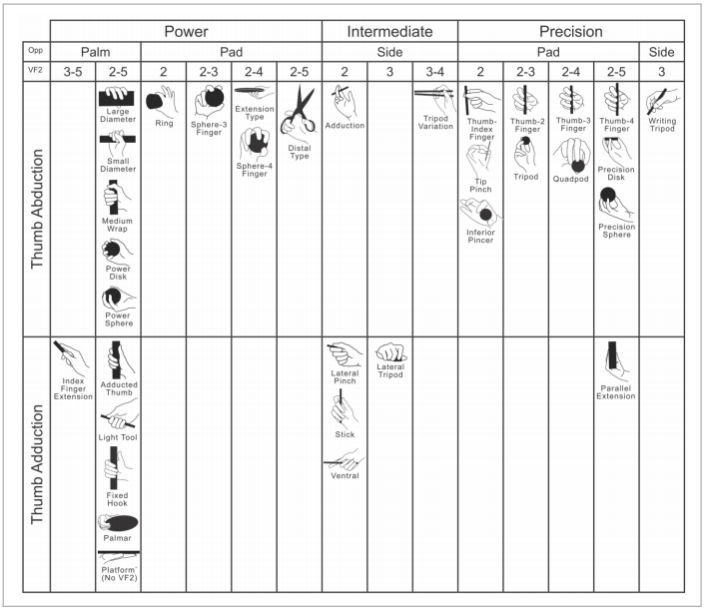
This revision highlighted the importance of the interplay between the index, thumb and middle finger in the human hand that form grasps such as the tripod grasp (MacKenzie and Iberall, 1994), and lead to presenting new grasp types, that are variations of this interplay with an additional unit (finger) such as: external precision grasp or writing grasp (Patkin, 1981), dynamic tripod (Parry, 1966), tripod grasp variation and adduction grasp (Kamakura et al., 1980) (see Fig-

ure 3.6 [page 36]). These grasps are called dynamic grasps (Kapandji, 1974), and represent the hand when it can still act while grasping an object, such as writing using a pen or cutting with scissors.

Cutkosky and Howe (1990); Cutkosky and Wright (1986) also focused on power requirements in their taxonomy, and extended the work of Napier (1956). They further added 9 and 7 sub-grasps in the power and precision grasps categories respectively as shown in Figure 3.7a [page 37].



(a) Cutkosky and Howe (1990) classified grasp types based on power and precision requirements. Image courtesy of Cutkosky and Howe (1990)



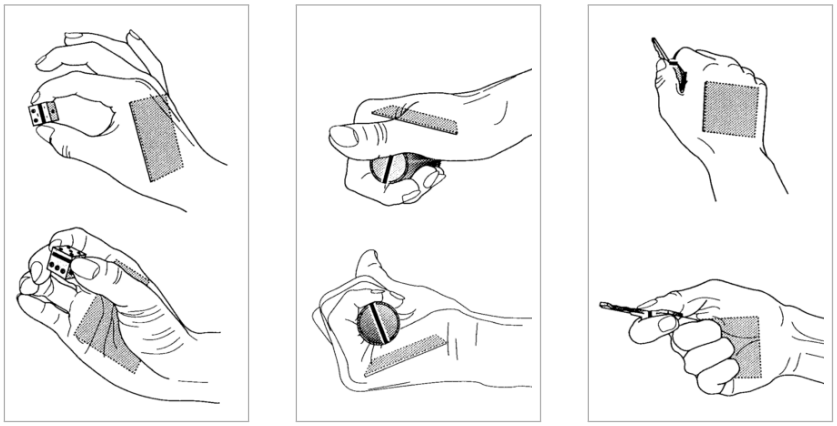
(b) Latest comprehensive grasp taxonomy, that is labelled by the authors (Feix et al., 2014) as the “most complete in existence”. Image courtesy of Feix et al. (2009)

Figure 3.7: Full descriptive taxonomies in current research for physical grasping

Their classification was based on power and precision grasp attributes, where they emphasised

on security and stability (ability to resist slipping) for power grasps, and dexterity and sensitivity (accuracy of the fingers in carrying large motions and sensing small changes in position and force) for precision grasps. In addition, they also used object properties (shape and size) to further refine their detailed taxonomy (MacKenzie and Iberall, 1994).

Iberall et al. (1986) presented another way to classify grasp types by focusing on the fact that at least two forces are applied in opposition to each other against the surface of the object being grasped, and they used the term "opposition" to describe three different directions along which the hand can apply forces (see Figure 3.8 [page 38]), where a grasp can then be formed using combinations of these directions: pad opposition (direction of hand surfaces are parallel to the plam as shown in Figure 3.8a [page 38]), palm opposition (direction of hand surfaces is perpendicular to the plam as shown in Figure 3.8b [page 38]) and the side opposition (direction of hand surfaces is transverse to the palm as shown in Figure 3.8c [page 38]).



(a) Pad Opposition (b) Palm Opposition (c) Side Opposition

Figure 3.8: The three ways a hand provides opposition around objects, where combinations of these ways form different grasp postures. Solid black lines show the opposition vectors in the object, and the shaded area shows the plane of the palm. Images courtesy of MacKenzie and Iberall (1994)

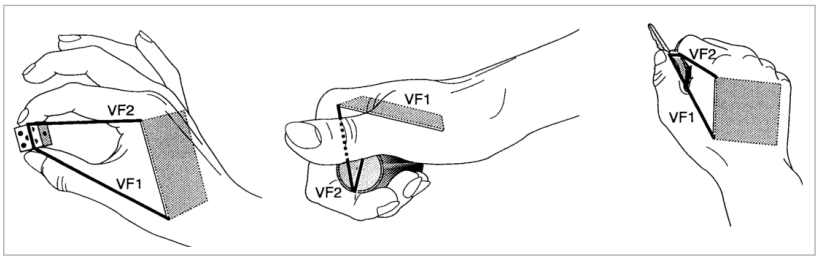


Figure 3.9: Oppositions described in terms of virtual fingers. Direction of virtual fingers is parallel to the plam in the Pad opposition, perpendicular to the plam in the Palm opposition and transverse to the palm in the Side opposition. Image courtesy of MacKenzie and Iberall (1994)

Arbib et al. (1985) have also presented the virtual finger (VF), which they defined as an abstract representation for a group of fingers and hand surfaces applying an oppositional force on the object being grasped (see Figure 3.9 [page 38]). For example, if you grasp a teacup, you will usually be able to fit 2 fingers in the inside of the handle of the cup (index and middle), and the thumb will be pressing on the outside of the handle. In this case, the thumb will be VF1, the two fingers in the inside of the handle will be grouped into VF2, VF1 and VF2 will be applying forces opposed to each other, and the remaining fingers that are pressed outside of the mug handle (fourth and fifth) will be VF3 as they counteract any task related torque if the mug is rotated towards the hand.

More recently, Feix et al. (2014, 2009, 2016) presented a comprehensive taxonomy that is based on aforementioned taxonomies, where they recorded two housekeepers and two mechanics in their professional working environment for around 8 hours. Grasp types in the footage of participants were then classified according to power/precision requirements, opposition type (pad, palm, side) where virtual fingers are also defined and thumb status (adducted/abducted). Combining these classification methods, in addition to using the thumb status as a grasp choice parameter, is unique to this taxonomy (Feix et al., 2009). Their final taxonomy resulted in 33 grasps, that could be further reduced to 17 if merging grasp types into one standardised grasp is feasible (see Figure 3.7b [page 37]).

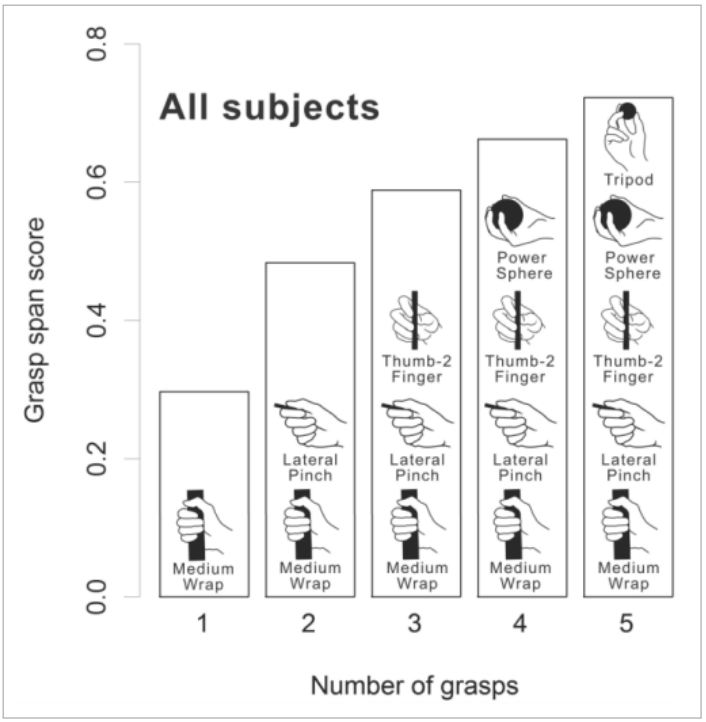


Figure 3.10: The most versatile grasp types for the majority of graspable objects. A higher grasp span means a higher versatility for a given grasp type. Image courtesy of Bullock et al. (2013)

Using the same methodology, Bullock et al. (2013) further analysed this taxonomy to further segment a set of grasps that can be used for the majority of graspable objects. This work introduced

the grasp span, a novel metric that is used to assess the versatility of a grasp to handle different object types. This metric was applied to over 19 hours of recorded footage and more than 9000 grasps, and presented 5 grasp types that this work deemed as the most versatile grasps that can be chosen for the majority of objects, namely: Tripod, Power Sphere, Thumb-2 Finger, Lateral Pinch and Medium Wrap (see Figure 3.10 [page 39]).

3.4.2 Objective Quantification

Taxonomies are considered a descriptive methodology to classify grasps, and MacKenzie and Iberall (1994) argues that a deeper understanding of the human grasp can be achieved through quantitative approaches.

Early work of Jacobson and Sperling (1976) presented a detailed coding system to quantitatively describe hand postures in healthy and injured hands. This work used film analysis, where grasp types were labelled in terms of hand surfaces involved and their positions, finger joint angles, contact surfaces of the fingers and palm with the object and the relationship between the hand and the object. Even though this approach was criticised for being time-consuming and unreliable (MacKenzie and Iberall, 1994), this coding system managed to form a taxonomy involving 8 grasp types Sollerman (1980).

Cutkosky and Howe (1990); Cutkosky and Wright (1986) presented an expert system called "GRASP-Exp" for choosing suitable grasps through observing mechanists working with parts and tools. The expert system asked the mechanists questions regarding the task requirements (dexterity and power) and object properties (size and shape), a posture was then chosen by the expert system through a hierarchical analysis of the task requirements (input) that were mapped to a suitable posture (output).

Iberall et al. (1988) developed a simulated neural network to choose an opposition (pad, palm or side) for a set of task requirements. The task in this work is presented as an input layer, where task requirements such as surface length, object dimensions, force magnitude, and precision are taken into account, and the trained neural network then chooses a suitable opposition (output) by weighting activation values of the input task requirements. In contrast to expert systems (Cutkosky and Howe, 1990; Cutkosky and Wright, 1986), this technique of encoding postures does not require for inputs (task requirements) and outputs (oppositions) to be explicit, instead the network learns the mapping process. Uno et al. (1993) also presented a neural network approach to determine optimal grasp types, where the network developed went under two phases, the learning phase and the optimisation phase. In the learning phase, the first input to the network consisted of visual images of objects that are different in shape and size (cylinders, spheres and prisms were used with sizes varying from 3cm to 7cm), and the second input was hand postures that were acquired using a data glove that is equipped with 16 sensors, where two hand postures were used, pad opposition and palm opposition. The network was trained on the relationship between the chosen objects and hand postures through repeated grasping in a trial and error manner, and a criterion function was then used in the optimisation phase to form an

optimal hand posture for the given object. While this model integrates visual and motor information in grasp choice, the limited number of inputs and outputs may not be representative of all dexterous human grasps (MacKenzie and Iberall, 1994).

Ekvall and Kragic (2005) explored programming a robot that is capable of grasping using humans that demonstrate the natural physical grasps first. Using magnetic trackers mounted on a data glove to monitor hand/finger (index, thumb and small) movement and orientation, this work evaluated 10 grasp types and three methods of grasp classification: finger position based, arm movement trajectory based and a hybrid system combining the latter two methods. Their work found that finger-based classification outperforms arm trajectory movement classification, and their hybrid system was the best method performance wise, as it overcame some of the shortcomings in grasp detection presented by the other two methods (finger bases and arm movement based. Although this work presented valuable insights into automated grasp classification, the most accurate presented (hybrid) could only detect around 70% of the grasps investigated. A similar approach in grasp classification using programming by demonstration using virtual grasping was implemented by Aleotti and Caselli (2006) in a virtual environment. Even though only 11 grasp types were assessed in this work, this approach using wearable sensors and tactile feedback showed promising grasp detection results for the grasps included in their analysis (82.8% to 94%).

More recently, Cai et al. (2017) presented an egocentric vision based grasp classification in unstructured environments. This method implemented state of the art computer vision techniques to detect hand features from video data that is recorded using a wearable camera. Hand features are then extracted to encode hand appearance and motion during the interaction, where their developed grasp classifiers are trained to distinguish between different grasps that are based on the taxonomy of Feix et al. (2009). Grasp classifiers finally quantify the visual similarities between the extracted features and the grasps of the used taxonomy. This approach showed promising results with reported 92% accuracy in grasp recognition in a laboratory setting. However, using this approach in clustered real world environments is still problematic, with the reported drop in grasp recognition accuracy to 59%. Liu et al. (2017) investigated grasp recognition and manipulations using EMG signals and multi-sensory information. This work utilised EMG (muscles), force (fingers) and tracking (hand) sensors to quickly segment 5 power grasping actions and 10 manipulation tasks, and even though this work does not address object properties in grasping actions, it still showed up to 92% accuracy in grasp detection.

3.5 Grasp Planning

3.5.1 Grasp Constraints

Dexterous hand and reach to grasp movements are designed to find integrated and continuous solutions to the neurophysiological and biomechanical constraints of any posture or move-

ment (Castiello, 2005). The motor system follows various constraints, especially in executing an action or posture such as a grasp, or a hand movement such as a reach to grasp movement where automation and rapidity are required (Jeannerod, 2006). Different constraints affect a reach to grasp movement in its different phases (MacKenzie and Iberall, 1994), for example, if you reach to grasp a glass, your hand will be shaped in accordance with object properties (size and weight if grasping a real object) due to anatomical (e.g. number of fingers), biomechanical (e.g. joints and degrees of freedom) and task (e.g. object properties and time limit) constraints that dictate how a grasp is formed.

Taking into account and analysing grasp constraints is important in choosing the optimal grasp (León et al., 2014), and is also essential in evaluating the quality of a grasping posture or movement. Iberall and MacKenzie (1990) presented a summary of the sources of constraints that impact grasp choice and motion (see Table 3.1 [page 43]), where constraints are divided into three main groups: high level constraints (e.g. mood and test time limits), physical constraints (e.g. object dimensions and arm reach) and sensorimotor constraints (e.g. finger pads and hand structure). This summary builds on previous studies (MacKenzie and Marteniuk, 1985; Marteniuk et al., 1987) that addressed grasp constraints, and is advantageous in assessing the complex interaction of object properties, environmental attributes and the anatomical and experience of users in grasping interactions.

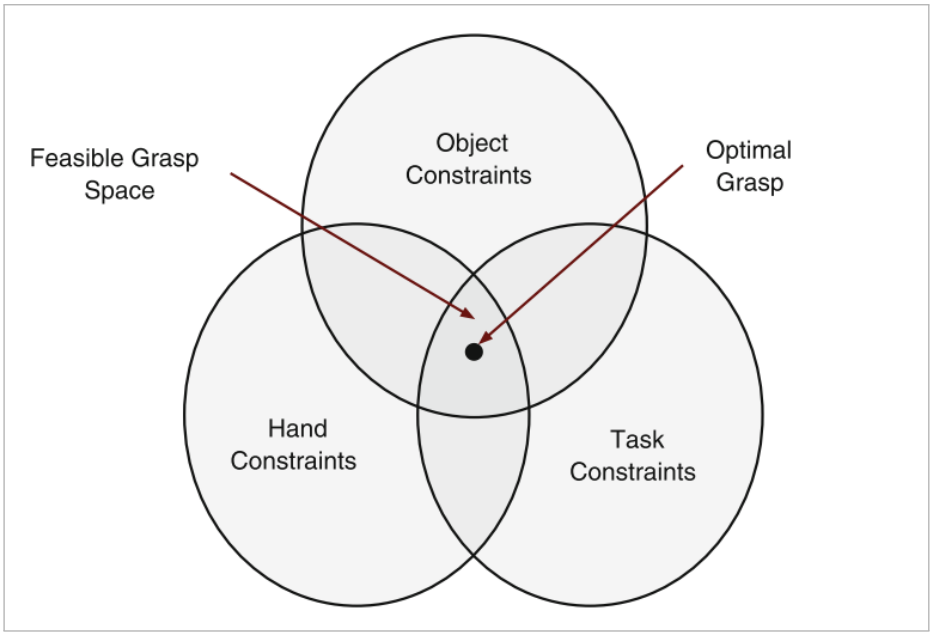


Figure 3.11: The three main grasp constraints that should be met in order to perform or choose an optimal grasp type. Image courtesy of León et al. (2014)

The work in this thesis is focused on analysing grasping virtual objects, and many of these constraints are not valid or present as they are proposed for grasping real physical objects. Constraints that are present in the framework of analysing grasping virtual objects in this work are addressed, and whether these constraints play the same role in grasping virtual objects as they

Table 3.1: Summary of the different types of grasp constraints. Table adapted from Iberall and MacKenzie (1990)

Group	Constraint Type	Examples
High Level	Social/Cultural	• don't stick out elbows • stick out little finger
	Motivational	• thirst • anger
	Informational	• convey affection • anger
	Functional	• don't drop object • manipulate object • move as quickly and as accurately as possible
Physical	Object Properties	• intrinsic (texture, surface length, weight, etc) • extrinsic (location, distance, environment, etc)
	Biomechanical/Mechanical	• kinematics • dynamics • limitations on force generation due to bones, muscles, tendons, ligaments, skin • effect and use of pads
Sensorimotor	Neural	• temporal and spatial limitations on CNS • pyramidal tract needed for fractionated finger movements • sensory info needed to sustain movement • sensory information needed to preshape hand • tonic vibration reflex
	Perceptual	• types, locations, and response characteristics of receptors • numerous tactile receptors in pulps with small receptive fields
	Anatomical/Physiological	• structural limitations on movements, directions, and extents • length of phalanges • additional muscles in index and little finger • pads • anatomical couplings
	Evolutionary/Developmental	• evolutionary pressures; five fingers • pyramidal tract develops in about eighth month

do for real physical objects is investigated. For this the key guideline provided by Napier (1956) is followed, that states that the grasp choice and action must satisfy imposed object, task and hand constraints, where power and precision features of the hand can match these requirements (see Figure 3.11 [page 42]).

3.5.2 Grasp Phases

Early work of Woodworth (1899); Jeannerod (1984) first described reaching to grasp or goal direct movements as a two-phased motion, an initial non-uniform or ungoverned motion, followed by a controlled final adjustment. In the first phase, the hand moves towards the object where fingers are preshaped in preparation for a grasp. In the second phase, any errors (e.g.

wrong postures and size misestimation) that occurred during the first phase are corrected, and the fingers are wrapped around the object. Jeannerod (1981) also suggested that the first phase in a grasp is faster than the second contacting and corrective stage. The literature expanded these phase to include task requirements such as manipulation, translation and release (see Figure 3.12 [page 44]). Grasp phases are differently named in various domains, such as: preload phase → loading phase → lifting phase (Johansson and Westling, 1984), set → preshape → enclose → hold → release (Ro et al., 2000; Debowy et al., 2001) or initialisation → approach → execution (Rijpkema and Girard, 1991). Even though the names of grasp phases vary, three main movements can be distinguished: reaching, grasping and manipulating/translating (León et al., 2014).

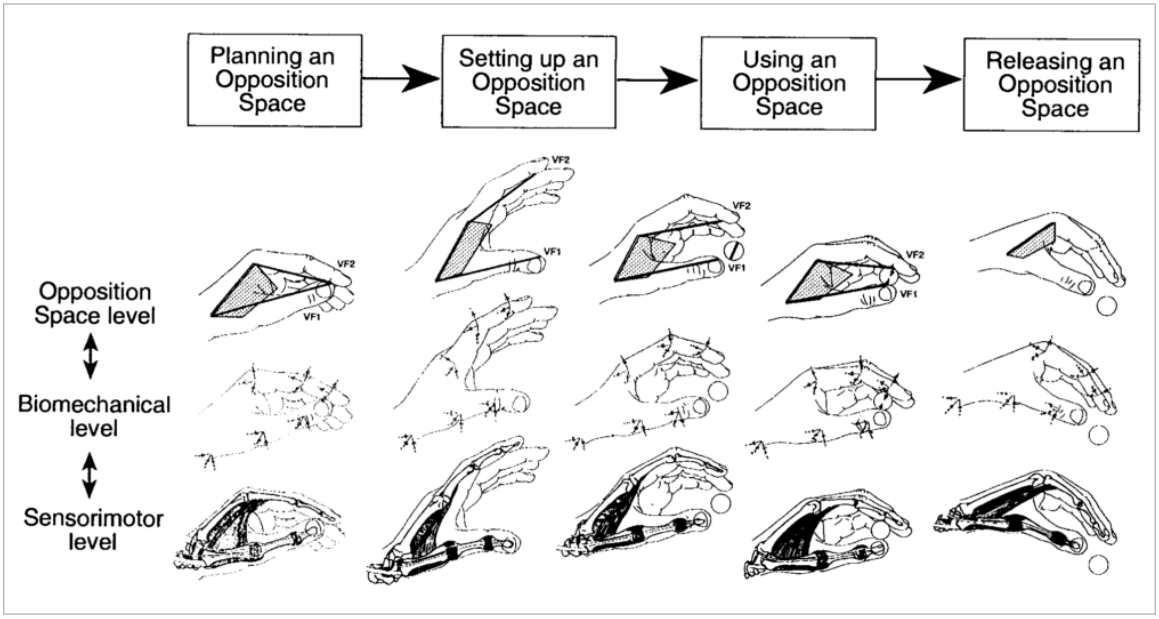


Figure 3.12: Grasp phases: a grasp starts from a resting posture in the pre motor planning stage (Opposition Space Planning), followed by opening of the hand in preparation for contact with the object (Setting Up Opposition Space). Manipulation or translation of the object then occurs once contact is made with the object (Using Opposition Space), and then finally the opposition space (i.e. object) is released. During these phases, motor commands from the CNS are generated at the Opposition Space, Biomechanical and Sensorimotor levels for the movement to satisfy all the constraints in a given task. Image courtesy of MacKenzie and Iberall (1994)

Grasp phases are summarised in full by Gordon (1994); Jeannerod (1986); MacKenzie and Iberall (1994) as:

3.5.2.1 Planning

Anderson (1985) suggested that visual perception can be divided into an initial stage where objects and shapes are extracted from a scene and a later stage where shapes and objects are recognised. Planning for a grasp starts before any movement is made, and that starts with perceiv-

ing the intrinsic and extrinsic properties of the object to be grasped when only visual feedback and perception are available. Intrinsic properties are internal properties of the object such as shape, weight, texture, hardness and size. Intrinsic properties can be perceived visually or haptically, and they affect grasp choice, namely shape and size, as they constrain the opposition type and position of hand placement. Extrinsic properties are spatial properties of the object to be grasped in an egocentric space such as location, orientation and distance. Similar to the intrinsic properties in this phase of a grasp, extrinsic properties are also perceived visually, and can also impact a grasp choice and planning. Perceiving object properties in this phase is followed by choosing a grasp strategy and hand location/orientation. This is mainly implemented by the CNS that plans the best grasp strategy taking into account perceived object properties, task constraints and personal experience.

3.5.2.2 Reaching

Trajectory planning is mediated through a combination of neurophysiological and mechanical functions. Once a grasp strategy is chosen, this goal of grasping an object is then transformed by the CNS to a motor action that involves joint angels and muscle activity. After planning for a grasp, the arm moves towards the object, and the fingers are preshaped to accommodate the size and shape of the object.

3.5.2.3 Pre-load

In this phase, the fingers press on the object obeying task constraints where the wrist and arm provide support to overcome any external forces (if grasping real objects), and this provides a stable grasp. It is also noted that the grip force increases in this stage when grasping a real physical object, and also provides additional feedback that is haptic. In goal or task oriented reach to grasp movements, this additional sensory feedback corrects any errors in grasp choice, position or strategy.

3.5.2.4 Load

This phase can be seen as a preparation stage for the goal of a reach to grasp movements. After the object is stably grasped and the grasp is corrected (if necessary), depending on the goal of the reach to grasp movement (translating or lifting) there will be a parallel increase in the load and grip forces (Bennett and Castiello, 1994). Meaning that the load force applied by the grasp becomes the gravitational force on the object (e.g. lifting a glass off a table).

3.5.2.5 Transition

Using biomechanical arm and wrist support and complex CNS commands, the object is translated in this phase from an initial position to a target location in a stable manner. The nature of the translation is task dependent (task constraint), thus this phase can also be holding an object

in a static state in the air, known as a static phase, where load and grip forces and object position are constant (e.g. holding a glass).

3.5.2.6 Release

Also known as an unloading phase, this phase occurs once the grip and load forces decrease in parallel. Releasing an object can also occur if the grasping posture is not compliant with shape and object constraints. For example, grasping an 8cm in width object using a 10cm grasp aperture that can cause the object to be released if virtual or fall/slip if real.

3.6 Summary

This chapter reviewed the theoretical definitions and parameters of the action of physical grasping, and detailed the proposed concept of freehand grasping of virtual objects in exocentric AR. Understanding the parameters and classifications of physical grasping is important prior to translating and analysing this form of physical interaction to AR environments. Current research in analysing physical grasping presents a comprehensive analysis the different parameters of grasping actions, these parameters are largely ignored in current AR applications and need to be taken into account in order to analyse this form of interaction, as they can potentially influence grasping performance.

Chapter 2 showed that grasping is widely used in different AR UIs (see Section 2.4.1 [page 14]), however this chapter highlighted that various elements of reach to grasping movements such as grasp phases, types and constraints are not addressed in current literature, or are falsely labelled (e.g. grasp types are inaccurately labelled). This can be problematic for AR systems that require physical grasping to be used in certain applications such as medical or engineering tasks, thus understanding the physical parameters of grasping is essential to the evaluation of the proposed concept of freehand grasping of virtual objects. In addition, freehand grasping will enable analysing the accuracy of this form of physical grasping in AR by taking into consideration physical grasping parameters.

In the following chapter, the methods and novel metrics used to analyse the accuracy and usability of freehand grasping in exocentric AR are discussed, providing an overview of the four user studies in this thesis along with a description of the tools and environment developed for this analysis.

Chapter 4

Freehand Grasp Evaluation Methodology

4.1 Introduction

Measuring user interaction performance in AR environments is key to assessing the usability of a certain application, and also aids in identifying problems and limitations of the environment developed and/or the interaction technique used within the environment. In addition, measuring interaction performance can show human behaviour trends during the interaction that can potentially help in developing more usable AR applications. Methods for measuring human performance in current literature are categorised by Dünser et al. (2008); Helander (2014); Dünser and Billinghurst (2011) as: objective measurements (e.g. completion time and error rate), subjective measurements (e.g. questionnaires and user ratings), qualitative analysis (e.g. user observations and formal interviews), usability evaluation (e.g. heuristic evaluation and task analysis) and informal evaluations (e.g informal user observations and informal interviews). Methods for quantifying user interaction performance in current research fall within these categories, however measuring the accuracy of grasping and understanding the perceptual nature and potential problems of freehand grasping of virtual objects has not yet been explored.

In Chapter 2 different AR applications that implement and analyse grasping interactions were discussed. Current AR applications mainly assess grasping using the grasp aperture metric, that is the distance between the index finger and thumb. Grasp aperture is widely used for physical grasping to quantify user performance as it provides information regarding the hand opening and the spatial position of a grasp in relation to a certain object given the inherit haptic feedback provided by the hand during physical grasping. However, for natural user interfaces (NUI) where haptic feedback is not feasible to implement, grasp aperture does not provide all the required information regarding grasp accuracy due to the lack of haptic response. In particular, the spatial position of a grasp in relation to virtual objects in NUIs cannot currently be mea-

sured using only grasp aperture. New metrics alongside grasp aperture are therefore required to assess this form of interaction in NUIs.
This chapter presents the proposed methods in this work to measure the accuracy of freehand grasping in exocentric AR (Section 4.2 [page 48]). The proposed methods are used in four independent user studies to analyse grasp accuracy and address the problems found in exocentric AR:

- **Study 1 (Chapter 5)** is the baseline study in this work that aims to measure user grasp performance in exocentric AR
- **Study 2 (Chapter 6)** then measures user grasp performance in exocentric AR using dual view feedback to address the key problems found in Study 1 using this proposed form of visual feedback
- **Study 3 (Chapter 7)** measures user grasp performance in exocentric AR using drop shadows to assist users in locating virtual objects by providing an additional visual cue during interaction
- **Study 4 (Chapter 8)** finally measures user grasp performance in exocentric AR using user grasp tolerances that are based on the grasping data collected in Study 1. This study explores how freehand grasping can still be usable without requiring users to be highly accurate in their grasping performance

In this chapter, the baseline environment and methods used for the four studies will be outlined, and the commonalities and changes to the baseline methods used for the different studies will be discussed. Data collected from these four studies are used to analyse user grasp performance, and to address the different problems and limitations of freehand grasping (see Figure 4.11 [page 64]). Section 4.2.1 [page 48] first presents the grasp parameters chosen for this body of work that are based on the grasp parameters for physical grasping discussed in Chapter 3. In Section 4.2.2 [page 50] the grasp model and novel metrics proposed in this body of work to measure the accuracy of freehand grasping in exocentric AR are defined. Section 4.2.3 [page 54] then outlines the AR environment developed and details the different components of the environment. Finally Section 4.2.4 [page 62] describes the experiment protocol adopted for all the studies in this thesis.

4.2 Methodology

4.2.1 Grasp Parameters

4.2.1.1 Grasp Taxonomy

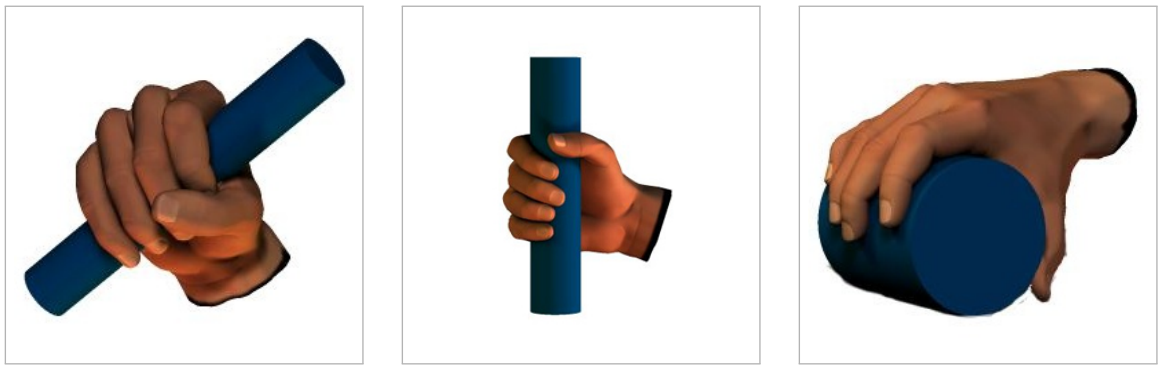
Grasp taxonomies are prominently present in the literature of grasping real physical objects (see Section 3.4.1 [page 34]), yet a grasp taxonomy for grasping interactions in AR environments is

still not available to use or adapt for this work, especially for the freehand form of grasping. Despite valuable previous gesture taxonomies presented by Piumsomboon et al. (2013); Wobbrock et al. (2009) that involved grasping for marker-based AR interactions, these taxonomies are highly influenced by the applications they were developed for, and grasping parameters such as grasp aperture, grasp constraints and grasp type are not addressed. In addition, the majority of gestures outlined in these taxonomies are bi-manual interactions due to the tabletop environment used, thus are not classified as grasps by definition.

4.2.1.2 Grasp Type

For this work one grasp type is assessed, that is the medium wrap grasp (see Figure 4.2a [page 50]), defined as the most common manual human grasp (Bullock et al., 2013; Feix et al., 2014) (see Figure 3.10 [page 39]). Their work is based on the grasp taxonomy of Feix et al. (2009) for physical grasps that is known to be the most complete grasp taxonomy to date (Feix et al., 2016)(see Figure 3.7b [page 37]). This work only focuses on this one grasp type as a control measure to the first studies looking into freehand grasping accuracy in this thesis. In addition, this grasp is suitable for the object types and experiment conditions assessed in this work. Findings from this work can later be validated for other widely used grasp types.

The medium wrap grasp is classified as a power grasp, and is an intermediate grasp between the small diameter(see Figure 4.1a [page 49]) and large diameter grasps (see Figure 4.1c [page 49]). These three grasps are also known as “full hand wrap” grasps and are relatively similar (Feix et al., 2016), thus one can easily interchange their grasp type between these three grasp types depending on the power and precision requirements of a given task (see Figure 4.1 [page 49]), however users in all the studies in this thesis are explicitly instructed to recreate the medium wrap grasp only for all objects and conditions.



(a) Small Diameter (b) Medium Wrap (c) Large Diameter

Figure 4.1: Full Hand Wrap Grasps

4.2.1.3 Grasp Phases

This body of work is mainly focused on reaching to grasp virtual objects in different phases and the phases defined by Gordon (1994); Jeannerod (1986); MacKenzie and Iberall (1994) are adapted. This work focuses on the reaching, pre-load and transition phases, and the planning, load and release phases are not included in the analysis. This adaptation of phases is justified by the essential information missing in interacting with virtual objects in an exocentric AR environment.

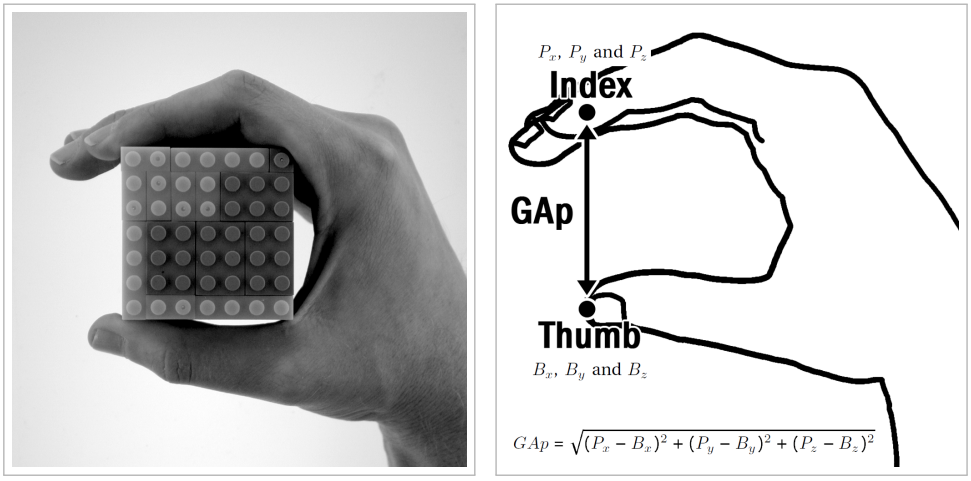
The planning phase is largely dependent on neurological activity in the human brain, that is outside the scope of this work. The load phase requires object weight in order to occur as a phase, this object property is missing in the proposed exocentric AR system, and even with the addition of haptic feedback, this phase would still not be valid to analyse due to the missing weight of virtual objects. Finally, the release phase is not included due to the findings found in Chapter 8 [page 172], where users showed inconsistency in grasp choice and aperture over a specified interaction distance, thus enabling releasing or dropping objects during a movement is not feasible.

4.2.2 Grasp Metrics

Designing new metrics that quantify the accuracy and nature of freehand grasping of virtual objects using the medium wrap grasp was required for this work, as an evaluation of this kind of interaction in an AR context has not yet been explored.

4.2.2.1 Grasp Aperture

Edsinger and Kemp (2007) define grasp aperture as the distance between the thumb and the index finger, and it is a common metric in human manipulation studies.



(a) Medium Wrap grasp

(b) Grasp Aperture (GAp)

Figure 4.2: 4.2a Grasp type (Medium Wrap) analysed in this work. 4.2b Grasp Aperture (GAp) used for quantifying grasp accuracy

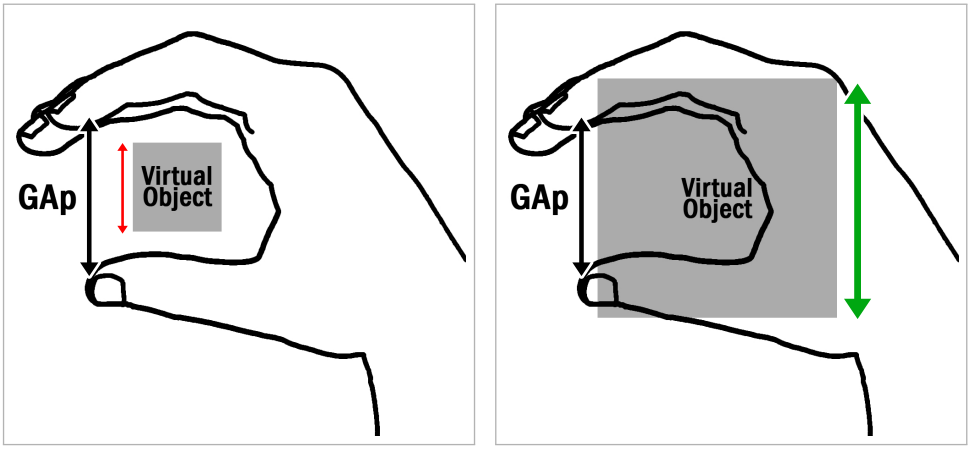
To measure how accurately users estimate the size of the virtual object the aperture of a user's grasp is applied, based on the work of Edsinger and Kemp (2007). Grasp Aperture (GAp) is defined in Equation 4.4 [page 51] to be the distance between a users thumb tip and index finger tip (see Figure 4.2b [page 50]). Grasp Aperture (GAp) is given as

$$Xaxis = \sqrt{(P_x - B_x)^2} \quad (4.1) \qquad Yaxis = \sqrt{(P_y - B_y)^2} \quad (4.2) \qquad Zaxis = \sqrt{(P_z - B_z)^2} \quad (4.3)$$

$$GAp = \sqrt{(P_x - B_x)^2 + (P_y - B_y)^2 + (P_z - B_z)^2} \quad (4.4)$$

Where P_x, P_y and P_z are the co-ordinates of the index finger tip, and B_x, B_y and B_z are co-ordinates of the thumb tip. These measurements are taken from the Kinect sensor used in this work, where x is measured from the centre of the sensor, y from ground and z from sensor (see Section 4.2.3.3 [page 58]).

GAp will provide information regarding how accurately users the size of the virtual object presented, and will show whether users overestimate the size of the virtual object presented (i.e. have a GAp bigger than the object size) (see Figure 4.3a [page 51]) or underestimate it, that is essentially when the grasp penetrates the virtual objects (i.e. perform a GAp smaller than the object size) (see Figure 4.3b [page 51]).



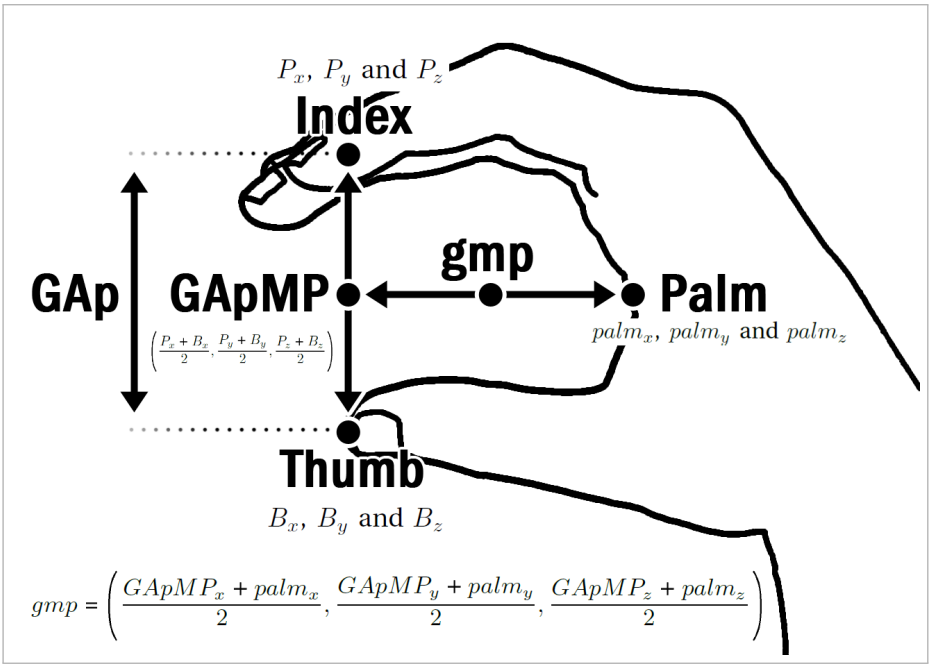
(a) GAp Overestimation (b) GAp Underestimation

Figure 4.3: 4.3a Overestimation and 4.3b underestimation of virtual objects size measured using GAp

4.2.2.2 Grasp Displacement

Sensory information in the point of contact with objects is an important component of a grasp strategy, this component is missing in freehand grasping due to the lack of tactile feedback. Thus solely depending on GAp to asses freehand grasping may result in false analysis, as GAp does not provide information about the spatial placement of a grasp.

A grasp displacement metric is proposed in this work that provides spatial information about a grasp in conjunction with a reference point in space. Similar to GAp, grasp displacement



(a) Grasp Middle Point (gmp)

Figure 4.4: gmp that is used for measuring Grasp Displacement ($GDisp$) in the x ($GDisp_x$), y ($GDisp_y$) and z ($GDisp_z$) axes

coordinates are taken from the Kinect sensor used in this work, where x is measured from the centre of the sensor, y from ground and z from sensor. The middle point of the grasp aperture ($GApMP$) is first calculated as

$$GApMP = \left(\frac{P_x + B_x}{2}, \frac{P_y + B_y}{2}, \frac{P_z + B_z}{2}\right) \tag{4.5}$$

Where spatial information about the index and thumb fingers are used to calculate the middle point of the grasp aperture in x, y and z axes. Only using $GApMP$ is insufficient, as it provides information about the position of a grasp, but not how the position of a grasp compares to the position of a virtual object, and even if a grasp is placed on an object, $GApMP$ will provide an offset that will not provide an accurate reflection of the grasp displacement. In addition, as a grasp requires users to estimate both the size of the virtual object and the spatial position, the grasp aperture (GAp) would also not be a suitable measure if used alone. Therefore to measure the position accuracy of both the user's hands against the virtual object a measure of the grasp middle point (gmp) is defined. Here gmp is defined in Equation 4.6 [page 52] as the position in the grasp relating middle point between the grasp aperture middle point $GApMP$ and the users palm (see Figure 4.4a [page 52]). The grasp middle point (gmp) is calculated as

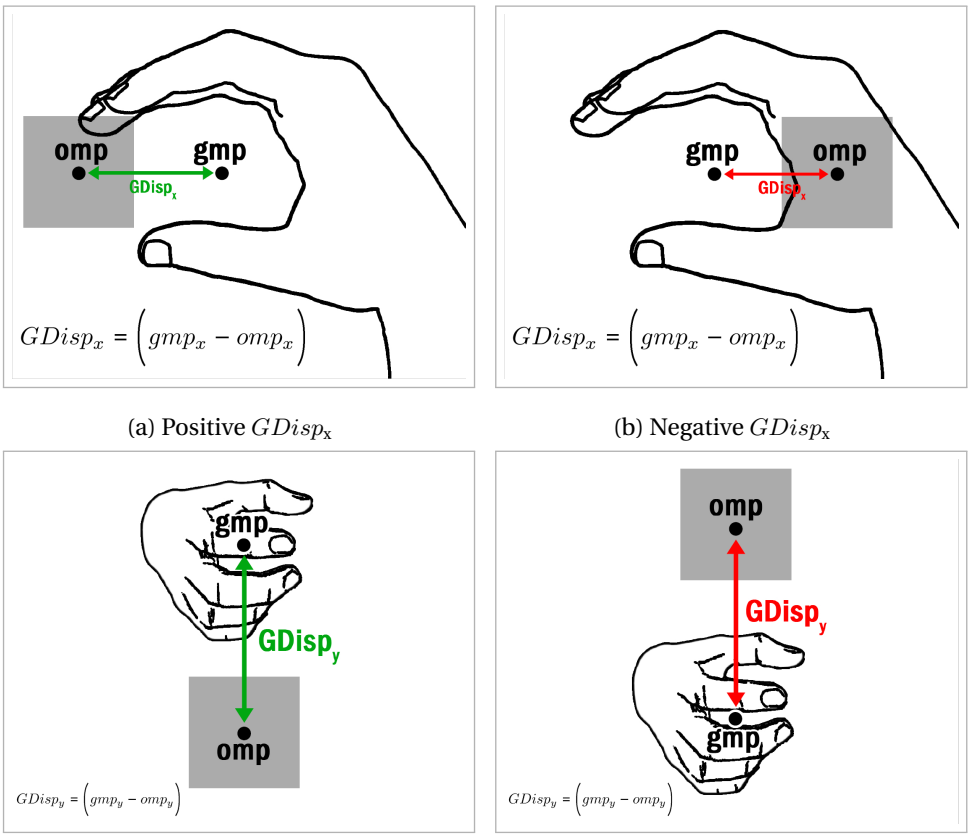
$$gmp = \left(\frac{GApMP_x + palm_x}{2}, \frac{GApMP_y + palm_y}{2}, \frac{GApMP_z + palm_z}{2}\right) \tag{4.6}$$

Where $palm_x$, $palm_y$ and $palm_z$ are the positions of the palm. Using the placement of gmp, Grasp Displacement ($GDisp$) is then calculated by subtracting the position of the middle point

of a virtual object (omp) from the gmp. This results in the distance from the middle point of the grasp to the middle point of the virtual object in the x ($GDisp_x$), y ($GDisp_y$) and z ($GDisp_z$) axes (see Equation 4.7 [page 53]).

$$GDisp_{xyz} = \left(\left(gmp_x - omp_x\right), \left(gmp_y - omp_y\right), \left(gmp_z - omp_z\right)\right) \tag{4.7}$$

$GDisp$ will provide information regarding the spatial position of the grasp in relation to the virtual object presented in the x ($GDisp_x$), y ($GDisp_y$) and z ($GDisp_z$) axes. $GDisp$ will also show the direction in which users place their grasp with relation to the virtual object in all axes, this can potentially provide insights regarding user behaviour and preferences during freehand grasping. The direction of grasp displacement in the x, y and z axes (i.e. whether positive or negative) is determined by the positioning of the grasp (i.e. gmp) in relation to the centre of the physical Infra-red (IR) sensor on the Kinect that is the origin of its coordinate system (i.e. $x = 0$, y, $z = 0$). In this coordinate system and from the user's point of view, x grows to the right of the sensor, y grows up and z grows out in the direction the sensor is facing.



(a) Positive $GDisp_x$

(b) Negative $GDisp_x$

(c) Positive $GDisp_y$

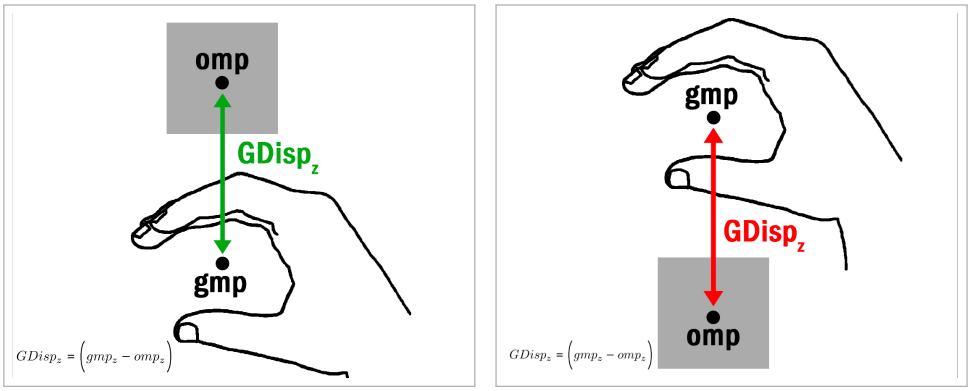
(d) Negative $GDisp_y$

Figure 4.5: Direction of gmp placement in relation to the omp in the x (4.5a and 4.5b) and y (4.5c and 4.5d) axes measured using $GDisp_x$ and $GDisp_y$

For $GDisp_x$, a positive displacement shows the Grasp Middle Point (gmp) is further placed to the right than the Object Middle Point (omp) (see Figure 4.5a [page 53]). In contrast a negative

$GDisp_x$ shows the gmp is further placed to the left than the omp (see Figure 4.5b [page 53]). For $GDisp_y$, a positive displacement shows the gmp is at a position higher than the omp (see Figure 4.5c [page 53]), whereas a negative displacement shows the gmp is at a position lower than the omp (see Figure 4.5d [page 53]).

For $GDisp_z$ that is essentially the depth estimation of a virtual object, the terms underestimation and overestimation are opposite to those of depth perception. Thus in this work, depth refers to the distance from the feedback monitor and not the user as in depth perception studies. Overestimation refers to the placement of the gmp at a further point away from the sensor than the omp, and this results in a positive $GDisp_z$ (see Figure 4.6a [page 54]). Underestimation refers to the placement of the gmp at a closer point to the sensor than the omp, and this results in a negative $GDisp_z$ (see Figure 4.6b [page 54]).



(a) Depth Overestimation (b) Depth Underestimation

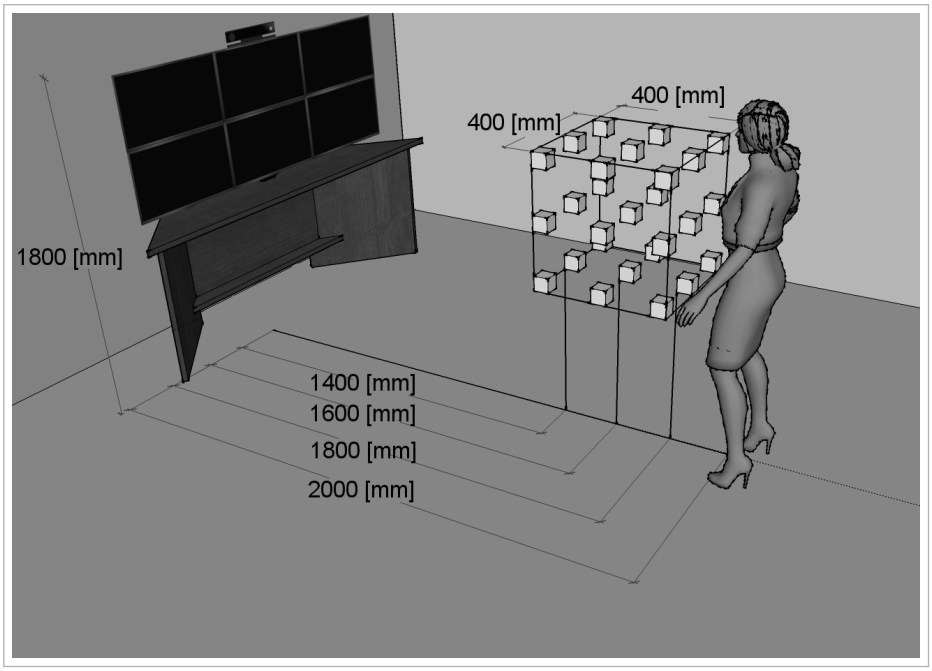
Figure 4.6: 4.6a Overestimation and 4.6b underestimation of the virtual object's position in the z axis measured using $GDisp_z$

4.2.3 Baseline Environment

4.2.3.1 Setup Overview

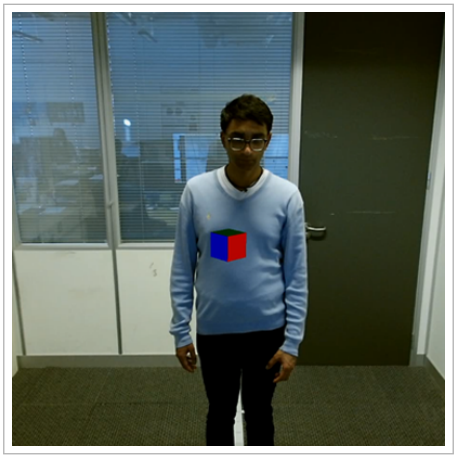
A concise description of the environment developed in this work is an interactive exocentric mixed reality environment, where users can directly interact with the virtual objects presented that are viewed indirectly by users (through a monitor) in an exocentric manner (users looking at the environment from the outside). The baseline exocentric AR environment developed integrated the use of a Microsoft Kinect 2, a (HD) video camera, and a SyncMasterX6 feedback monitor. Participants stood 2000mm away from the feedback monitor (size: 62in × 30in, resolution: 5760 × 2160), displaying a composited real-time mirrored scene overlaying virtual objects with the video feed. Grasping parameters (GAp, $GDisp$) are measured from the sensor, not to test biomechanics of the hand but to quantify errors in spatial positioning and aperture estimation.

Across all the studies in this thesis, the physical configuration of the system strictly and consis-



(a) Experiment Setup

(b) User View



(c) Sensor View

Figure 4.7: Setup of the system developed

tently followed the recommendations of Kinect's V2 manufacturers[1] to ensure ideal operating conditions of the sensor. Accordingly participants stood 2000mm away from the sensor under controlled and constant lighting conditions, the sensor was placed at a height of 1800mm and tilted at an angle of $13.78°$ to show the full working space around participants and to eliminate any significant self-occlusion problems (see Figure 4.7 [page 55]). The test coordinator was seated at a room corner behind the feedback monitor that is outside the field of view of the sensor. Users were instructed not to move during their grasping interaction, and in order to ensure this a mark was placed on the floor for users to stand on, alongside a box that was taped to the floor and acted as a barrier in front of users to avoid any movement.

[1] http://support.xbox.com/en-GB/xbox-360/kinect/kinect-sensor-setup

To analyse the impact of the spatial position of virtual objects on freehand grasping in terms of the $GDisp$ and GAp measures proposed in this work, the methodology of Stockmeier et al. (2003) is adapted. Their method assessed grasping virtual objects in a seated tabletop setting using haptic sensors on the index finger and thumb, where they placed projected virtual objects on a mirror in 27 different positions in all axis (x, y and z). Following this work, virtual objects are positioned in 27 different positions in all three axis in a NUI setting (where each object is displayed once per position)(see Figure 4.7a [page 55]), covering a range of 400mm from participants that is within the mean biomechanical arm reach of participants (See Chapter 5 [page 66] for detailed positions). While this method can be limited as it does not take into account user performance in every possible position in the interaction space or potential user movements during a grasping movement that is outside the scope of this work, adapting this arrangement method of virtual objects in 27 different positions allows analysing user performance and accuracy in all three axes that is the main aim for this first analysis of grasping accuracy in AR, and can potentially provide valuable insights regarding user performance in different interaction (e.g right, left, centre, top, bottom) and reaching (e.g. close to body, far from body) regions. In addition, this method can also be easily adaptable depending on the biomechanical features of users.

To analyse the impact of object size of virtual objects on freehand grasping in terms of the $GDisp$ and GAp measures proposed in this work, users are presented with cubes and spheres in 6 different sizes that are: 40mm, 50mm, 60mm, 70mm, 80mm and 100mm in one position that is the centre position shown in Figure 4.7a [page 55]. This range of sizes is chosen based on the guidelines outlined by Feix et al. (2014) for grasping real objects, where they illustrated that the hand is rarely challenged to perform a grasp aperture equal to or larger than 100mm, and is most comfortable performing a grasp aperture that is less than 50mm. In compliance with these guidelines, object sizes would range from 40mm to 80mm, with an addition of the 100mm size in order to test the applicability of this real world grasp aperture range in freehand grasping in AR environments.

This baseline environment is used in all the user studies in this thesis. Minor modifications to this environment are made in the second, third, fourth and fifth studies to assess different aspects of freehand grasping, and to address some of the potential problems using this interaction technique (see Figure 4.11 [page 64].

4.2.3.2 System Architecture

The exocentric AR environment in this work was developed in a controlled laboratory environment using the following tools:

- **Kinect's V2 SDK**: provided real-time depth and skeletal information of the scene and hand joints in particular, namely the thumb, index finger palm and wrist. This information was extracted from the motion capture device utilised (Kinect) and implemented using C++ programming language. This information was used to facilitate freehand grasping and

enable layer based occlusion to allow free movement behind, in front or simultaneously in front and behind virtual objects

- **Autodesk Maya**[2]: was used for modelling 3D objects, namely cubes and spheres in different sizes and orientations. Maya provided two main files for each model, namely ".OBJ" that provides information regarding the structure of the object (e.g. position of vertices that make up polygons) and a ".MTL" file that provides information regarding the texturing or colouring of the modelled objects. Contents of these files can be manually edited if changes to the appearance or position of the object are required without the need of using Maya directly

- **OpenGL**[3]: open graphics library (OpenGL) was used for real-time reading, loading, texturing of the three dimensional (3D) objects modelled in Maya. OpenGL is considered to be an industry standard tool for developing interactive 3D graphics applications and is currently the most supported 3D graphics library (Liu and Wu, 2018), hence can easily be integrated in interactive systems such as the one developed in this work. In addition, the capabilities of OpenGL in drawing 3D objects extend beyond the basic and abstract rendering features used in this work, thus making it a suitable tool to use in any potential future routes that stem from this work concerning the impact of different rendering effects on grasping performance. In this work OpenGL loads the ".OBJ" and ".MTL" files, scans through the files to extract structure and texturing information and finally draws the loaded virtual object using this information in its own independent scene (an OpenGL scene, in this case, is essentially a window showing the object on a fully black background, and is displayed on a separate window from the main video scene that is coming from the Kinect's camera). This process is mediated by an OpenGL virtual object loader (i.e. script), and such loaders are widely available for developers to use or modify based on the needs of the application being developed owing to the wide use of OpenGL in interactive 3D graphics applications

- **OpenCV**[4]: open computer vision library (OpenCV) was used to merge or overlay rendered virtual objects on the video scene coming from the Kinect's camera. This was implemented by masking out the virtual object displayed in the OpenGL scene, which was then overlayed on the video scene in its corresponding position. In addition, OpenCV was also used to add feedback information on the video displayed for users (such as test number and countdown timer) and to extract the real world spatial position of virtual objects using image-based computer vision techniques. Even though OpenCV can be used independently to track hand movements using vision based techniques (Pham et al., 2018), these techniques are still often prone to unreliable tracking due to light changes, increasing dis-

[2] http://www.autodesk.com/products/maya/overview
[3] http://www.opengl.org/
[4] http://opencv.org/

tances from a camera or hand articulations, for these reasons the Kinect sensors is used instead in this work to track hand and finger motions during grasping actions

4.2.3.3 Tracking

Tracking of users and particularly their hand joints is facilitated in this work by the Kinect V2 motion capture sensor. Use of Kinect in a similar exocentric environment to the one proposed in this work was previously validated by Hough et al. (2015) to measure user interaction performance during freehand bi-manual interaction with virtual objects in different sizes.
Kinect V2 is widely used to assess interaction and applications in AR environments (recent examples in Reither et al. (2018); Gavrilova et al. (2018)), and is suitable for this work as it is capable of accurately tracking essential joints for the grasp model presented in this work, namely the index finger, thumb and palm. In addition, the markerless based nature of the Kinect sensor allows users to perform freehand grasping in a NUI where they are not mounted with any wearable sensors and are capable of grasping in a way that resembles real world grasping.
Kinect V2 is capable of tracking subjects up to 6m away from the sensor (Gonzalez-Jorge et al., 2015) and also offers a wide FOV (up to 70°) that allows users to view the interaction space in full while performing freehand grasping. This wide tracking range is useful for freehand grasping as performing different grasping motions such as reaching and fine adjustments of fingers are not restricted by the tracking sensor and users are accordingly capable of freely moving around in the environment if needed. Kinect has a random error in depth measurement that depends on the distance away from the sensor, that can range from a few millimetres to 4cm at the maximum tracking range of the sensor, with the optimal distance for data acquisition being within 1-3 meters from the sensor (Khoshelham and Elberink, 2012). In order to mitigate the sensor's depth error, a repeated measures design has been used in all studies in this work, and the optimal working conditions for the sensor were strictly followed. The low cost, portable and non-intrusive nature of Kinect V2 sensors make it acceptable in the research community to provide interaction evaluations (Yang et al., 2015), especially for applications where wearable devices are not valid to use such as in medical applications (Lun and Zhao, 2018). Kinect V2 sensors also use the standard in motion capture that is Time of Flight (ToF) technology that is integrated in current state of the art hardware systems.

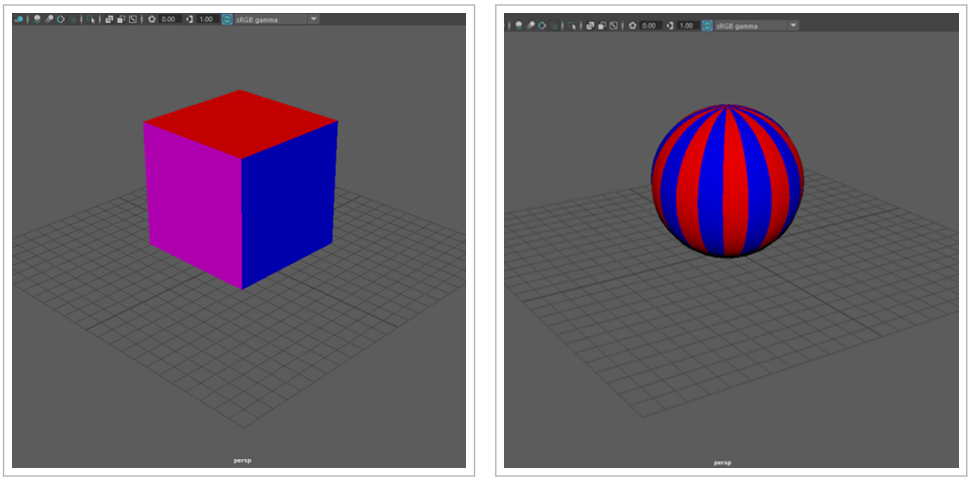
4.2.3.4 Virtual Objects

Cubes and spheres are the two virtual objects used to assess freehand grasping accuracy and usability in this work. Cubes and spheres are by definition "regularly shaped" objects with an equal distribution of mass (MacKenzie and Iberall, 1994), that have visible characteristic such as width, spatial density and radius of curvature which are utilised by the CNS during the process of grasp planning. This is particularly important in freehand grasping where haptic feedback is not available, as features such as spatial density (i.e. texture) and size can be perceived using only vision (Klatzky et al., 1987).

Grasp planning, matching object size using the hand opening in particular, is also largely dependent on object features that are the object width (influences grasp aperture size) and length (influences number of fingers used) (Newell et al., 1989), and the regularity of cubes and spheres object clearly offer two locations for grasping that are usually parallel to each other, and are known as opposable surfaces that result in an opposition vector with the magnitude being the size of the grasp aperture performed.

These opposable surfaces in cubes and spheres that are visually accessible are directly related to the grasp model presented in this work, as sizes of regular object (i.e. cubes and spheres) can be measured using grasp aperture (Jeannerod et al., 1990; Chan et al., 1990), unlike irregular or complex objects (e.g. mugs) where grasping a handle, for example, is not representative of the overall size of the object. Moreover, monitoring choice of opposable surfaces by participants and measuring the resulting opposition vector (i.e. grasp aperture) in this work will aid in quantifying the accuracy of freehand grasping, and understanding human behaviour in freehand grasping that is still unclear in current literature.

Cubes and spheres in this work are coloured using default Maya materials in order to provide a sense of depth to avoid objects looking flat on the feedback monitor used in this work which can potentially hinder depth perception (see Figure 4.8 [page 59]). For cubes, each side is overlayed with a different colour in order to ease the process of distinguishing each side of the model (see Figure 4.8a [page 59]), and similarly for spheres vertical polygons are grouped and coloured differently all around the sphere to avoid them looking like a flat circle (see Figure 4.8b [page 59]). Rendering of cubes and spheres in this study is basic without any additional rendering features such as enhanced lighting, textures, reflections (see Figure 4.8 [page 59]). Given that this body of work is the first to analyse freehand grasping in exocentric AR, additional rendering features are excluded in order to avoid any potential perceptual bias, thus grasping is analysed in this basic form of rendering that is essentially the default rendering settings in Maya, allowing these attributes to be studied independently.



(a) Cube Object in Maya (b) Sphere Object in Maya

Figure 4.8: Virtual objects in this work

4.2.3.5 Feedback

Owing to the nature of this work being a visual and gesture NUI, visual feedback is the only form of feedback provided to users in this work (see Figure 4.9 [page 60]). Visual feedback in AR environments is considered to be a conventional type of feedback that is widely used (Prattichizzo et al., 2012), and in this work visual feedback is provided using a standard large SyncMasterX6[5] monitor (size: 62in × 30in, resolution: 5760 × 2160) that provides users with a mirrored image of their interaction and the environment in real time (see Figure 4.9a [page 60]), similar to visual feedback provided in TV virtual studios for presenters (Hough et al., 2015)(see Figure 4.9b [page 60]). This exocentric setting, where users view the environment from the outside, is suitable to facilitate freehand grasping that resembles grasping in the real world, for this reason users are not provided with visual feedback using HMDs for example, in order to avoid limiting their movement that can potentially lead to unnatural and uninformative perceptions (Gibson, 1950, 1966, 2014).

(a) Visual feedback in this work

(b) Visual feedback for freehand interaction in virtual TV studios

Figure 4.9: Visual feedback in exocentric AR

[5] http://www.samsung.com/us/support/owners/product/MD230X6

Occlusion, a visual perception phenomenon that occurs if an object hides or partially hides another object from view (Epstein and Rogers, 1995), is also implemented as a depth cue for users in this work across all the four studies. Occlusion offers information regarding the depth ordering of objects in a certain environment thus allowing users to judge the relative nearness of an object in relation to their field of view. Depth cues comprise an important component in scene and interaction realism (Cutting, 1997), and occlusion is considered to be the strongest and most consistent depth cue in comparison to other depth cues such as size constancy, accommodation and shadows, that can be trusted at any given distance where visual perception holds true (Wade and Swanston, 2013). Previous work (Hough et al., 2015) illustrated that implementing authentic occlusion which resembles that experienced naturally in the real world for freehand interaction creates a more realistic environment for users and viewers, in addition to providing reliable depth cues to users regarding the virtual objects displayed. This is particularly important in this work as grasping is usually subject to occlusion in the real world, thus occlusion is implemented using the depth data from the Kinect sensor to allow users to view their grasp as real world grasp where their hand, or individual fingers, can be in front of, behind or simultaneously behind and in front of the virtual object as shown in Figure 4.10 [page 61].

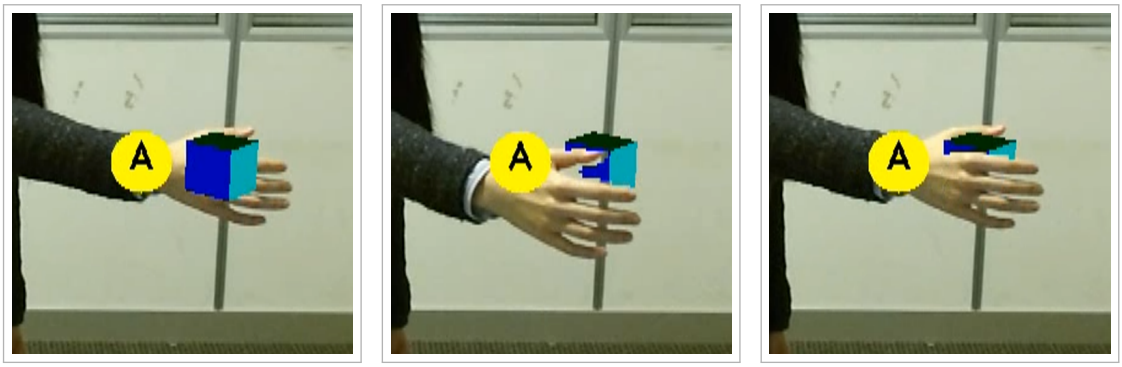


Figure 4.10: Occlusion handling in this work

In addition, implementing occlusion handling will also aid in analysing freehand grasping in terms of object position, particularly in the z axis where object depth can be challenging to interpret using only the feedback monitor, thus using occlusion handling is compatible with the overall environment in this work being a NUI without the use of additional wearable devices to aid in depth perception.

Use of the standard monitor as the device for visual feedback and occlusion as a depth cue is consistent throughout all the studies in this thesis. Minor additions to this baseline visual feedback are implemented in the second, third and fourth studies (see Figure 4.11 [page 64]).

4.2.4 Experiment Protocol

4.2.4.1 Participants

A sample size of 30 participants was chosen for each study in this thesis (different sample for each study). Participants volunteered to take part in the studies in this thesis from a population of university students and staff members. Participants were naive to the purposes of all the experiments in each of the four studies in this thesis, and their level of experience in AR systems ranged from novice to expert.
All participants were right-handed due to the fact that the majority of the human population is right handed (Oldfield, 1971), and the right side in humans tends to be generally stronger in human fetuses due to larger bone and muscle structures (Pande and Singh, 1971). The right hand is also favoured for interactions that require high forces (power grasps) due to the postural specialisation of the left hemisphere of the human brain, and this preference followed due to evolution for fine manipulation and precision grasping (Jeannerod et al., 1990). For these reasons, recruiting only right-handed participants is a control measure that is potentially reflective of the majority of the population in this first analysis of freehand grasping accuracy of virtual objects in AR.

4.2.4.2 Protocol

Before each study, participants completed a standardized consent form. Visual acuity was measured using a Snellen chart (where 1 is equivalent to 20/20 vision), and each participant was required to pass an Ishihara test to exclude for colour blindness. No participants suffering from colour blindness and/or with visual acuity of < 0.80 were included in the analysis.
Height, arm length and hand size of all participants were also measured prior to each study, this was done to ensure that aspects of the experimental design (such as object size and position) are within the biomechanical reach of participants. This was followed by initial training of the medium wrap grasp on real and virtual objects. The training session lasted for 5 minutes, during which users were trained on performing a medium wrap grasp on a virtual cylinder that was displayed on the feedback monitor and real physical objects that were essentially lego cubes in different sizes (40mm, 60mm and 80mm). A virtual cylindrical object was used in training instead of a cube or sphere to avoid any potential learning effects. Participants were not compensated in all the studies in this work, and all data collected was anonymised.

4.3 Studies and Hypotheses

This work consists of the four independent user studies that address different problems in freehand grasping, and also evaluate different methods to improve grasping performance and usability (see Figure 4.11 [page 64] for an overview of the four user studies in this work). The following is a summary of the primary aims and hypotheses that will be under test in the four

user studies:

- **Study 1: Measure user grasp performance in exocentric AR**
- **Hypotheses:**
  - **$H_{1.1}$:** Changes in object size do not have an effect on: a) grasp aperture and b) grasp displacement
  - **$H_{1.2}$:** Changes in object position do not have an effect on: a) grasp aperture and b) grasp displacement
- **Study 2: Measure user grasp performance in exocentric AR using dual view visual feedback**
- **Hypotheses:**
  - **$H_{2.1}$:** Dual view visual feedback in grasping virtual objects that change in size has no effect on grasp aperture and grasp displacement
  - **$H_{2.2}$:** Dual view visual feedback in grasping virtual objects that change in position has no effect on grasp aperture and grasp displacement
- **Study 3: Measure user grasp performance in exocentric AR using drop shadows**
- **Hypotheses:**
  - **$H_{3.1}$:** Adding drop shadows in freehand grasping of virtual objects that change in position has no effect on grasp aperture and grasp displacement
  - **$H_{3.2}$:** Adding drop shadows in freehand grasping of virtual objects that change in position has no effect on task completion time
- **Study 4: Measure user grasp performance in exocentric AR using user based grasp tolerances**
- **Hypotheses:**
  - **$H_{4.1}$:** Grasp tolerances (absolute and average) have no effect on task completion time and usability in grasping interactions

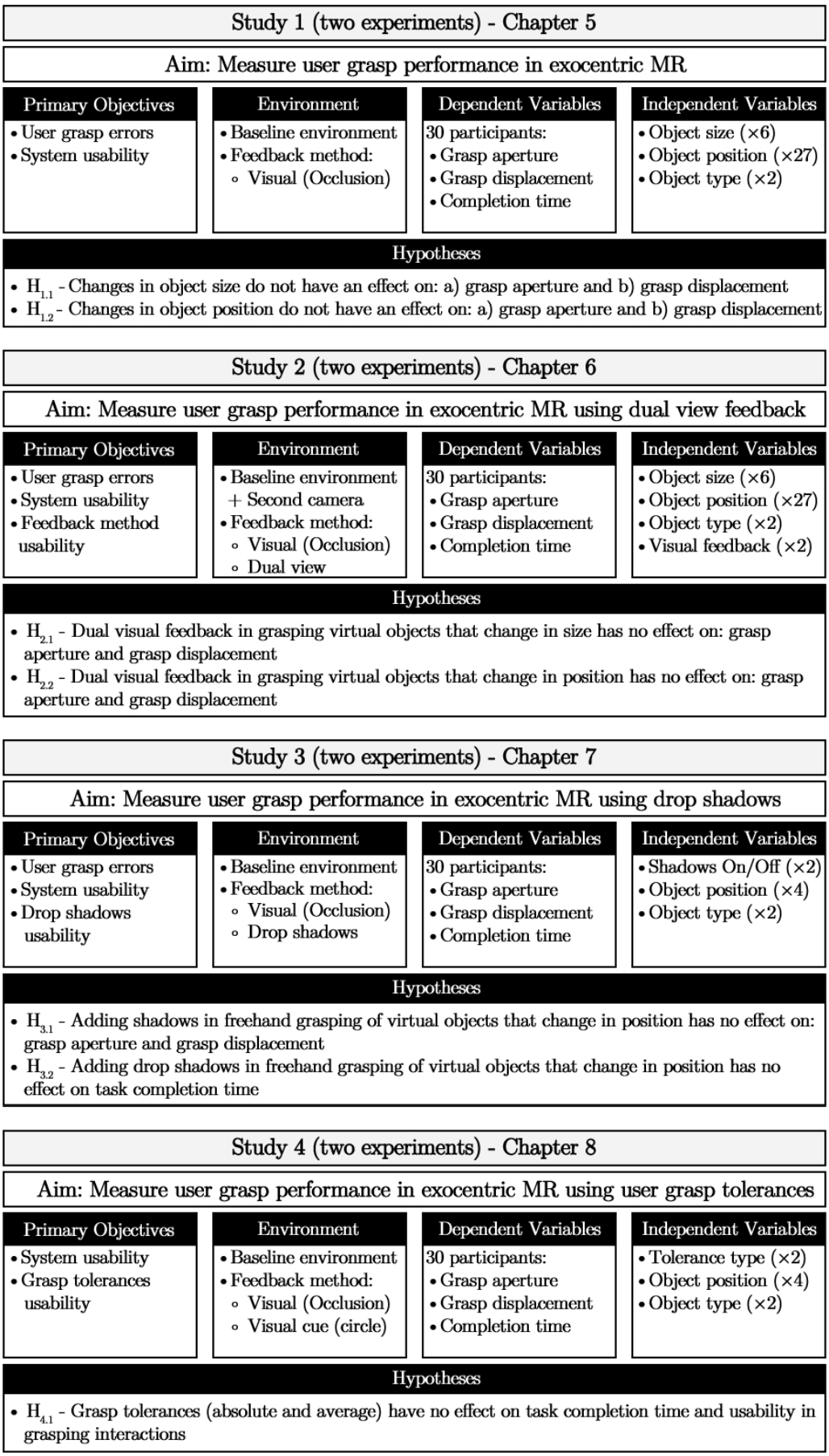


Figure 4.11: Structure of the four studies in this work

4.4 Summary

This chapter detailed the methods proposed in this work to measure freehand grasping accuracy in exocentric AR, and outlined the structure of the four studies presented in this thesis. The proposed methods discussed in this chapter will be used in all the following four user studies in this thesis.

Following the discussion of current research focused on measuring grasp performance for both real and virtual objects in Chapters 2 and 3, it was evident that there are currently no methods available to assess the accuracy of the proposed concept of freehand grasping in natural user interfaces. Current research mainly assesses grasp accuracy in AR environments using grasp aperture that is a method derived from grasping physical objects, with the aid of additional wearable based haptic feedback to compensate for the missing haptic feedback in AR environments. However this method is not feasible to use in natural user interfaces where the use of wearable sensors is not feasible, thus using grasp aperture only in the proposed work in this thesis would not be valid and will most likely result in false results. This is mainly due to the fact that grasp aperture provides information regarding the hand opening of a user, but not the spatial position of grasp in relation to a certain virtual object.

In this work grasp aperture (GAp) will be used alongside a novel proposed method that is grasp displacement ($GDisp$) that will provide the position of a grasp in relation to virtual objects in NUIs, by measuring the displacement from the grasp middle point (gmp) to the object middle point (omp) in three dimensional space. These methods will aid in measuring the accuracy of grasping against different virtual object sizes and positions in a natural user interface without the use of any additional wearable sensors.

The following chapter will present the collection and analysis of data in the first out of four user studies in this thesis, where GAp and $GDisp$ will be used to measure user grasping performance against different virtual object sizes, types and positions in the exocentric AR environment presented in this chapter. Findings from the study in the following chapter will highlight the key problems and limitations of freehand grasping that will be addressed in subsequent chapters.

Chapter 5

Study 1: Freehand Grasping Errors

This work was published in the proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) as "Analysis of Medium Wrap Freehand Virtual Object Grasping in Exocentric Mixed Reality" (Al-Kalbani et al. (2016*a*))

5.1 Introduction

Grasping is one of the primary forms of manual interaction between humans and the physical world. While this is the case, the dexterous versatility of the human grasp poses many challenges within virtual object interaction and as such, the objective quantification of these problems is largely unexplored. Subjective evaluation methods to assess interaction provide useful information regarding the ease and consistency of interaction or a certain interactive system, however, such methods do not provide enough information regarding interaction accuracy. For this reason, objective methods are instead used in current research to form a better understanding of human interaction accuracy and performance in AR environments. For example, Swan et al. (2015) measured depth judgement accuracy in matching and reaching interactions by measuring user distance from an ideal target location. Hough et al. (2015) also quantified the fidelity and plausibility of bi-manual interactions in a virtual studio AR environment using hand placement distance from ideal virtual object locations.

This chapter will present the first user study out of the four independent user studies in this work (see Figure 4.11 [page 64]) to assess the accuracy and usability of freehand grasping in exocentric AR. Using the baseline AR environment, grasp metrics and experiment protocol discussed in Chapter 4, this chapter will present an analysis of freehand grasping of virtual objects in two separate experiments, and illustrate the common errors within grasping virtual objects when they are presented within an exocentric AR scene displayed in front of the user. Section 5.2 [page 67] firstly outlines the design of the two experiments in this study in terms of the conditions under test, participants recruited, statistical model used and the experimental protocol. Section 5.3 [page 69] then discusses the data collected in the first experiment of this study, and

provides a comprehensive analysis of the results. This is followed by a discussion of the second experiment in this study and a comprehensive analysis of the results in Section 5.4 [page 77]. Finally Section 5.5 [page 96] provides the conclusions drawn from this study and a summary of the key outputs that will be further explored in following chapters.

5.2 Study Outline

5.2.1 Design

Two experiments were conducted in this study using the baseline setup detailed in Chapter 4 (see Figure 4.7a[page 55]):

- Experiment 1 to quantify the influence of object size and object type
- Experiment 2 to test the influence of object position and object type in x,y and z space

Conditions of both experiments are shown in Table 5.1[page 67], with the accuracy of a medium wrap grasp measured against the proposed metrics in this thesis; grasp aperture (GAp) and grasp displacement ($GDisp$). To represent the accuracy of a grasp independent of additional rendering, for both experiments, the baseline objects which have not undergone complex rendering and represent a simple abstract shape are used (see Section 4.2.3.4[page 58]).

Table 5.1: Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor

Experiment 1 - Object Size				
Condition	**Levels**			
Object Size [mm]	40 - 50 - 60 - 70 - 80 - 100			
Object Type	Cube and Sphere			
Experiment 2 - Object Position				
Condition	**Levels**			
Object Position (x, y) [mm]		Left	Centre	Right
	Top	-400, 1650	0, 1650	400, 1650
	Centre	-400, 1250	0, 1250	400, 1250
	Bottom	-400, 850	0, 850	400, 850
	** 9 positions were repeated in each z plane (1400mm - 1600mm - 1800mm), resulting in 27 positions in total*			
Object Type	Cube and Sphere			

Hypotheses

$H_{1.1}$: changes in object size do not have an effect on a) grasp aperture and b) grasp displacement (Experiment 1)
$H_{1.2}$: changes in object position do not have an effect on a) grasp aperture and b) grasp displacement (Experiment 2)

5.2.2 Participants

30 participants ranged in age from 19 to 62 (M = 30.43, SD = 9.78), in arm length from 480mm to 660mm (M = 552.40, SD = 43.80), in hand size from 160mm to 200mm (M = 186.80, SD = 10.40), in height from 1570mm to 1950mm (M = 1744.00, SD = 90.00) and 6 were female and 24 male. Taking into account balance in hand size, arm length, gender, age and height, participants were separated into two groups of 15 for each experiment. This separation of users into two groups was done to ensure that participants in the two groups are comparable and reflective of the population in terms of their physical features (i.e. arm length, hand size and height), and to avoid any bias in the results between the two user groups that may be influenced by the physical features of users.

5.2.3 Statistical Model

Statistical models used in the analysis of results in this study were validated using assumptions of different models. Kruskal Wallis H test (Kruskal and Wallis, 1952), a rank-based non-parametric test, is used to analyse the data collected in this study. Statistical significance of the Kruskal Wallis H test results is implemented using a post-hoc test for multiple comparisons using Dunn Test with Bonferroni correction (Dunn, 1961). This step is essential to check if there are any statistical differences between groups of independent variables.
In addition, Jonckheere-Terpstra (Terpstra, 1952; Jonckheere, 1954), a non-parametric test, is used to determine if there are any statistically significant trends between ordinal independent variables and continuous dependent variables (Field, 2012) (e.g. “increasing” or “decreasing” trend).

5.2.4 Protocol

This study followed the baseline experiment protocol outlined in Section 4.2.4.2[page 62] prior to collection of data.
All participants were instructed via a scripted description of the procedure of both experiments alongside written descriptions. The test coordinator explained the procedure again between each block of tests (i.e cube and sphere), and participants were allowed to rest before the presentation of every object. Each experiment was formed of a 5 minutes training/instructions session, 10 minutes of grasping a cuboid object, 5 minutes break and 10 minutes of grasping a spherical object (order of virtual objects counterbalanced).

5.3 Experiment 1 - Object Size

5.3.1 Experiment Design

A 2 × 6 within-subjects design was used, with two primary conditions: object size and object type. All 15 participants took part in both conditions. Every permutation of size × repetition for both object types was randomly presented to participants to exclude potential learning effects. In total, each participant completed 6 (sizes) × 5 (repetitions) × 2 (objects) = 60 trials and 900 grasps (60 trials × 15 participants). Each static grasp of every participant was recorded for 5 seconds (75 frames), this lead to collecting 67500 raw data points (900 grasps × 75 frames).

5.3.2 Procedure

For this first experiment, participants were instructed to accurately match their grasp aperture to the size and position of the object shown to them on the feedback monitor in the shortest time possible using a medium wrap grasp. Users were instructed to finish the task in the shortest time possible to test if the test conditions (i.e. virtual object size and type) have any impact on grasp accuracy.

Before interaction, an object (cube or sphere) appeared to participants on the feedback monitor, each object had 6 different sizes. Objects were positioned 1600mm away from the sensor, and 200mm away from participants, this position was unchanged throughout the experiment (see Figure 4.7a[page 55]). A countdown of 5 seconds followed by an auditory cue was used as an indicator for participants to start grasping the object.

During the interaction, all participants were instructed to verbally inform the test coordinator that they are satisfied with the grasp they have performed, and maintain the grasp for 5 seconds while the measurements are stored.

5.3.3 Results

Table 5.2: Significant Post-hoc Pairwise Comparisons - Experiment 1. Symbols are represented in constant order (● ○ ▲ △ ◆ ◇ ■ □) and represent significance in a post-hoc Dunn Test with Bonferroni correction using an α level of 0.01 for the following: ● GAp - Cube, ○ GAp - Sphere, ▲ $GDisp_{\mathrm{x}}$ - Cube, △ $GDisp_{\mathrm{x}}$ - Sphere, ◆ $GDisp_{\mathrm{y}}$ - Cube, ◇ $GDisp_{\mathrm{y}}$ - Sphere, ■ $GDisp_{\mathrm{z}}$ - Cube, □ $GDisp_{\mathrm{z}}$ - Sphere. No symbols indicate statistical similarity

Sizes	40	50	60	70	80	100
40		● ▲ ◆ ■	● ○ ▲ △ ◆ ■	● ○ ▲ △ ◆ ◇ ■	● ○ ▲ △ ◆ ◇ □	● ○ ▲ △ ◆ ◇ □
50			● ○ △ ◆ ◇	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇	● ○ ▲ △ ◆ ◇
60				● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇	● ○ ▲ △ ◆ ◇
70					▲ ■ □	● ○ ▲ △ ◆ ◇ ■ □
80						● ○ △ ◆ ◇
100						

5.3.3.1 Results - Grasp Aperture (GAp)

Statistically significant differences in GAp between different cube ($\chi^2(5) = 2824$, $p < 0.01$) and sphere ($\chi^2(5) = 1477$, $p < 0.01$) sizes were found. Significant adjusted post-hoc results are reported in Table 5.2 [page 69] (see ● for cubes, ○ for spheres).

A linear relationship is present between GAp and object size in the context of grasping real objects, thus the correlation between GAp and object size using a Jonckheere-Terpstra test for ordered alternatives shows a statistically significant trend of higher GAp with higher levels of cube size ($T_{JT} = 2.88 \times 10^8$, z = 53.58, $p < 0.01$) and sphere size ($T_{JT} = 2.75 \times 10^8$, z = 36.65, $p < 0.01$).

5.3.3.2 Analysis - Grasp Aperture (GAp)

Table 5.3: Descriptive Statistics of GAp, $GDisp_x$, $GDisp_y$, $GDisp_z$ and completion time for different sizes of cubes and spheres (Mean ± SD). For statistical significance ($p < 0.01$) between individual sizes, see Table 5.2 [page 69]

Object Type	Object Size	GAp	$GDisp_x$	$GDisp_y$	$GDisp_z$	Time
Cubes	40	66.31 ± 29.97	31.45 ± 14.10	-15.81 ± 12.15	-34.34 ± 65.58	4.28 ± 2.05
	50	73.89 ± 28.16	26.83 ± 13.33	-14.76 ± 11.93	-38.75 ± 60.73	4.43 ± 1.99
	60	76.18± 24.12	26.09 ± 14.18	-13.17 ± 12.57	-40.37 ± 62.10	4.16 ± 1.90
	70	80.38 ± 22.55	25.35 ± 14.75	-11.15 ± 10.37	-42.13 ± 55.54	4.25 ± 1.61
	80	80.13 ± 24.59	29.40 ± 14.69	-10.53 ± 10.62	-35.31 ± 56.68	3.96 ± 2.03
	100	88.77 ± 22.39	28.93 ± 12.31	-8.90 ± 12.36	-39.42 ± 68.15	4.48 ± 1.96
Spheres	40	65.73 ± 30.83	36.51 ± 13.31	-12.62 ± 12.76	-29.21 ± 60.92	3.89 ± 1.67
	50	63.71 ± 30.40	36.10 ± 13.91	-12.82 ± 13.03	-31.86 ± 62.98	3.57 ± 1.52
	60	66.48 ± 28.97	31.98 ± 14.96	-11.97 ± 11.90	-29.44 ± 57.61	3.45 ± 1.37
	70	72.18 ± 26.03	30.41 ± 14.52	-8.41 ± 12.08	-27.04 ± 61.96	3.33 ± 1.33
	80	70.16 ± 26.20	29.53 ± 13.59	-7.46 ± 12.89	-31.08 ± 62.04	3.23 ± 1.29
	100	77.24 ± 24.35	24.56 ± 14.26	-5.77 ± 10.39	-30.62 ± 57.20	3.41 ± 1.95

As shown in Table 5.3 [page 70], participants overestimated object size in grasping both cubes and spheres, where mean size overestimation occurred in 9 out of the 12 sizes for both objects under test across all users. Participants overestimated object sizes in all sizes up until the size that had the lowest mean difference between GAp and objects size (80mm for cubes (80.13mm ± 24.58), and 70mm for spheres (72.18 ± 26.03)). Users were comfortable in grasping these two sizes for both objects potentially due to the similarity of how these sizes were perceived by users to sizes of real world objects grasped on daily basis (e.g. water glasses). In contrast, participants were least accurate in matching their GAp to object size in the 40mm size for both cubes (66.31mm ± 29.97) and spheres (65.73mm ± 30.83), this is potentially due to the inconvenience of this small size where users were sometimes able to cover the whole object with their grasp, thus hindering their GAp estimation as they were not able in some cases to view the whole object due to its small size (see Figure 5.1 [page 71]).

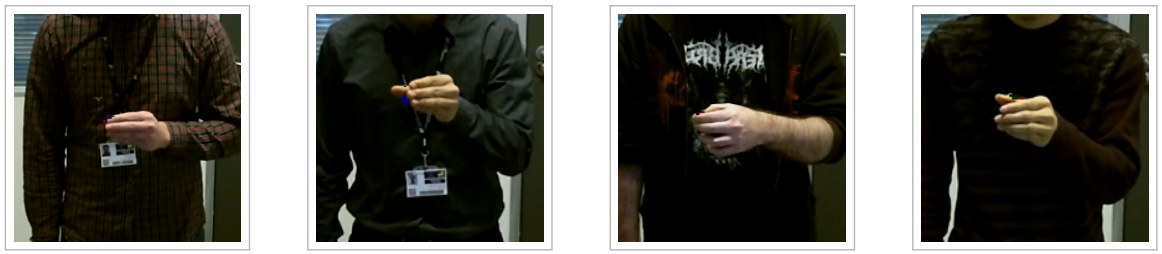


Figure 5.1: Users were able to fully occlude objects that were 40mm in size. This size showed the lowest accuracy in terms of matching GAp to object size

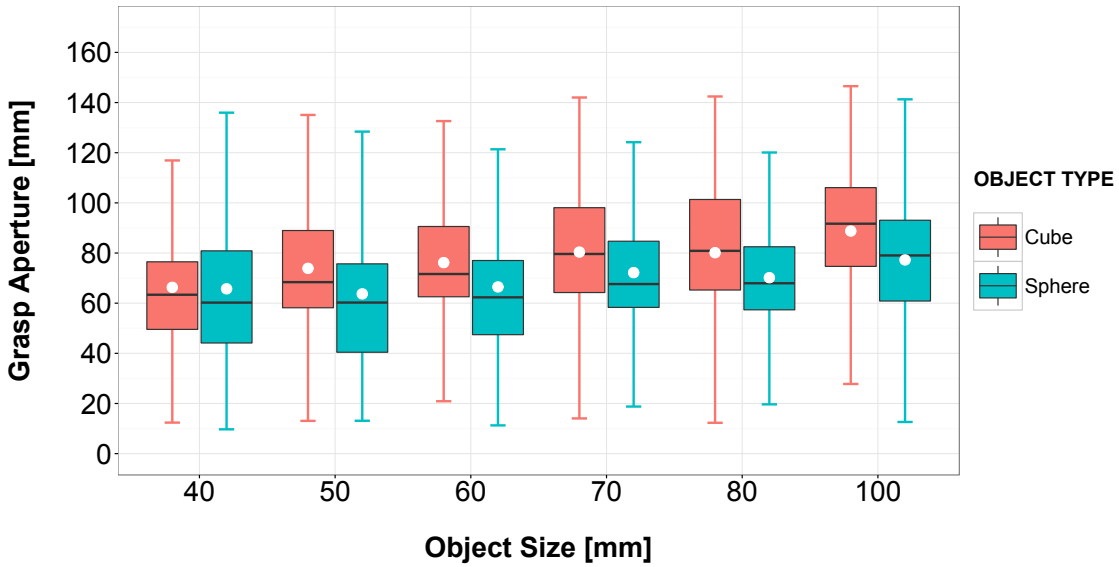


Figure 5.2: GAp for different object sizes in the 1600mm z plane. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

In addition, participants consistently underestimated the 100mm object size for both objects, with a mean underestimation of -11.23mm (SD = 22.39) for cubes and -22.76mm (SD = 24.35) for spheres. This potentially shows that users were grasping within a specific GAp, thus object size is underestimated once the size presented is outside of that working range of users.

Figure 5.2 [page 71] shows that the mean GAp ranged from 65.70mm to 88.80mm (SD = 20.48) across all sizes of both objects. This range shows that GAp in grasping virtual objects is between 65.70mm to 88.80mm regardless of object size or type. This shows that even though users, up to a point, increased their GAp with increasing object size, GAp is not proportional to object size in freehand AR grasping unlike grasping real objects.

5.3.3.3 Results - Completion Time

Statistically significant differences in completion time between different cube ($\chi^2(5) = 449$, $p < 0.01$) and sphere ($\chi^2(5) = 572$, $p < 0.01$) sizes were found.

5.3.3.4 Analysis - Completion Time

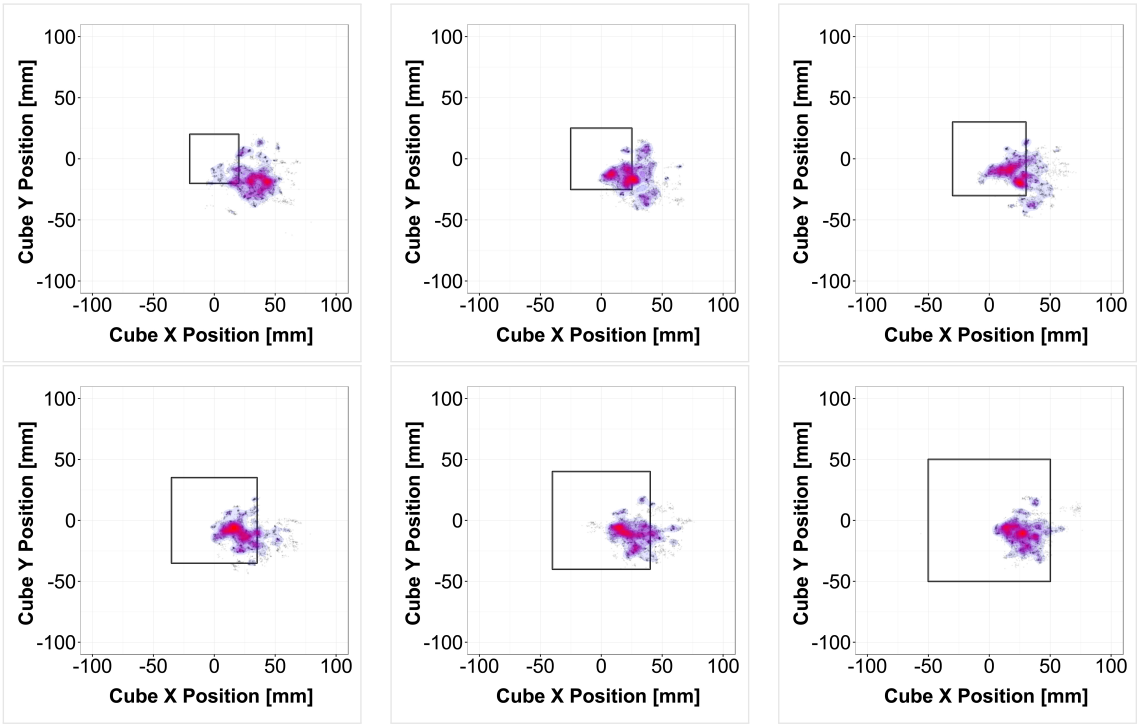
Mean completion time ranged from 3.23 to 4.48 seconds for both objects (SD = 1.30). Shortest completion time was present in the 80mm size for both objects (see Table 5.3 [page 70]), this could be an indication that the 80mm object size felt the most natural graspable size for participants. Even though statistically significant differences in completion time between different cube and sphere sizes were found, users showed similar completion time ranges for all object sizes and variation between different object sizes can be attributed to individual differences in perceptual tasks and the unchanged position of objects in this experiment (see Section 3.3.2 [page 32]), thus no significant trends between completion time and object size were found. Task completion times were initially expected to differ between different virtual object sizes, this was based on the guidelines outlined by Feix et al. (2014) for grasping real objects, where they illustrated that the hand is rarely challenged to perform a grasp aperture equal to or larger than 100mm, and is most comfortable performing a grasp aperture that is less than 50mm. However as shown in the results, these preferences for object sizes are not reflected by task completion times when grasping virtual objects.

5.3.3.5 Results - Grasp Displacement ($GDisp$)

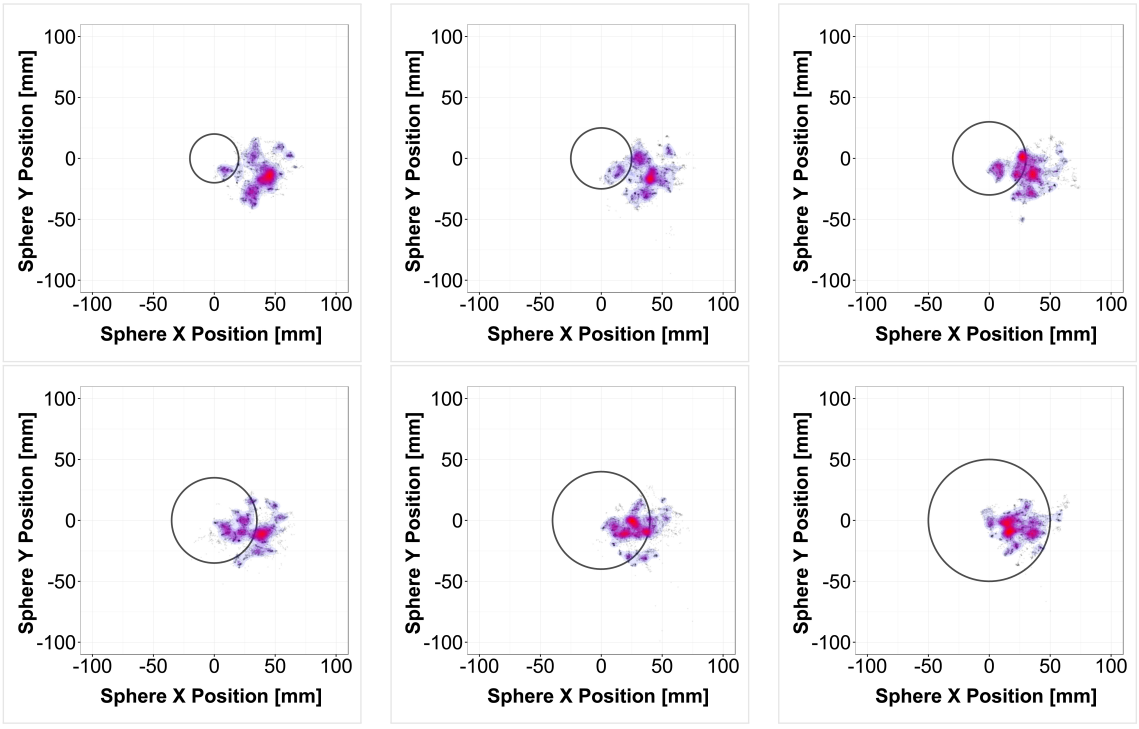
Statistically significant differences between different object sizes were found in $GDisp_{\text{x}}$ (cubes ($\chi^2(5) = 922$, $p < 0.01$) and spheres ($\chi^2(5) = 2728$, $p < 0.01$)), $GDisp_{\text{y}}$ (cubes ($\chi^2(5) = 1556$, $p < 0.01$) spheres ($\chi^2(5) = 1845$, $p < 0.01$)) and $GDisp_{\text{z}}$ (cubes ($\chi^2(5) = 135$, $p < 0.01$) and spheres ($\chi^2(5) = 82.77$, $p < 0.01$).
Full significant adjusted post-hoc results are reported in Table 5.2 [page 69].

5.3.3.6 Analysis - Grasp Displacement ($GDisp$)

$GDisp_{\text{x}}$
Participants had the highest mean $GDisp_{\text{x}}$ in the 70mm size for cubes (M = 42.13mm, SD = 55.54), and in the 50mm size for spheres (M = 31.86mm, SD = 62.98). Lowest mean $GDisp_{\text{x}}$ was present in the 40mm size for cubes (M = 34.34mm, SD = 65.58), and in the 70mm size (M = 27.04mm, SD = 61.96) for spheres.
Positive $GDisp_{\text{x}}$ was present for both objects in each size (see Table 5.3 [page 70]). This positive $GDisp_{\text{x}}$ is potentially due to the fact that users were interacting with virtual objects on the positive side of the sensor's coordinate system using their right hand (where x grows to the right of the sensor from the user's point of view). Across all sizes and participants, less $GDisp_{\text{x}}$ in grasping cubes (M = 28.01mm, SD = 14.08) than spheres (M = 31.52mm, SD = 14.67) was found, meaning more $GDisp_{\text{x}}$ was present for spheres as shown by the wider clusters in Figure 5.3b [page 73]. Participants also showed the highest mean $GDisp_{\text{x}}$ in the 40mm size for cubes (M = 31.45mm, SD = 14.10) and spheres (M = 36.51mm, SD = 13.31), the same size where participants showed the least accuracy in terms of matching GAp to object size.



(a) gmp placement along the x and y axes when grasping cubes



(b) gmp placement along the x and y axes when grasping spheres

Figure 5.3: gmp placement in the x and y axes of all participants in six virtual object sizes (in the order 40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 5.3a: black squares indicate cube sizes. 5.3b: black circles indicate sphere sizes. Density heat maps indicate gmp placement across participants (red indicates higher density)

Bounds of clusters presented in Figures 5.3a [page 73] and 5.3b [page 73] show similarity of Grasp Middle Point (gmp) placement for all cube and sphere sizes. Mean $GDisp_{\text{x}}$ ranged from 25.35mm to 31.45mm (SD = 12.31) across all cube sizes, and ranged from 24.56mm to 36.51mm (SD = 14.68mm) across all sphere sizes, showing a higher SD for spheres, as visualised by wider spread of clusters in Figure 5.3b [page 73]. However, SD differences within object sizes between cubes and spheres were comparable (see Table 5.3 [page 70]), indicating that contact of the gmp with the surface of the object was reflective of size growth of objects rather than movements by participants, this is evident by users showing the lowest mean $GDisp_{\text{x}}$ in bigger object sizes (70mm size for cubes (M = 25.34mm, SD = 14.75), 100mm size for spheres (M = 24.56mm, SD = 14.26)).

***GDisp*$_{\text{y}}$**

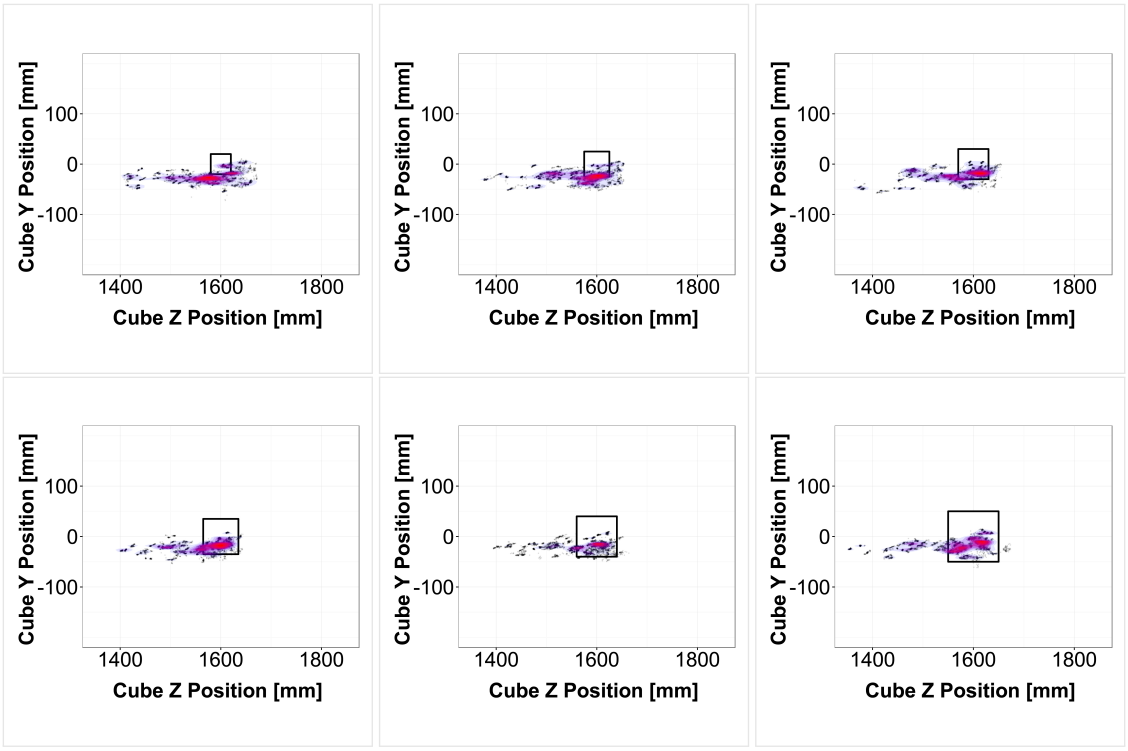
Negative $GDisp_{\text{y}}$ was present for both objects. This reveals that participants placed their gmp below the omp for both cube and sphere. Interestingly, participants chose a lower point to the omp and not a higher one, this is potentially attributed to participants trying to show parts of the objects presented to them on the feedback monitor. Participants also showed the highest mean $GDisp_{\text{y}}$ in the 40mm size for cubes (M = -15.81mm, SD = 12.15), and in the 50mm size (M = -12.82mm, SD = 13.03) for spheres, thus again potentially showing that small object sizes, where the grasp can cover the whole object, can be problematic in size estimation using GAp. Unlike bigger objects as users showed the lowest mean $GDisp_{\text{y}}$ in the largest 100mm object size for cubes (M = -8.90mm, SD = 12.36) and spheres (M = -5.77mm, SD = 10.39). Clusters in Figures 5.3a [page 73] and 5.3b [page 73] show that gmp placement was comparable for all cube and sphere sizes along the y axis. Mean $GDisp_{\text{y}}$ ranged from -8.90mm to -15.81 (SD = 11.94) in all cube sizes, and ranged from -5.77mm to -12.82mm (SD = 12.51mm) in all sphere sizes, showing a higher SD for spheres (see Table 5.3 [page 70]).

***GDisp*$_{\text{z}}$**

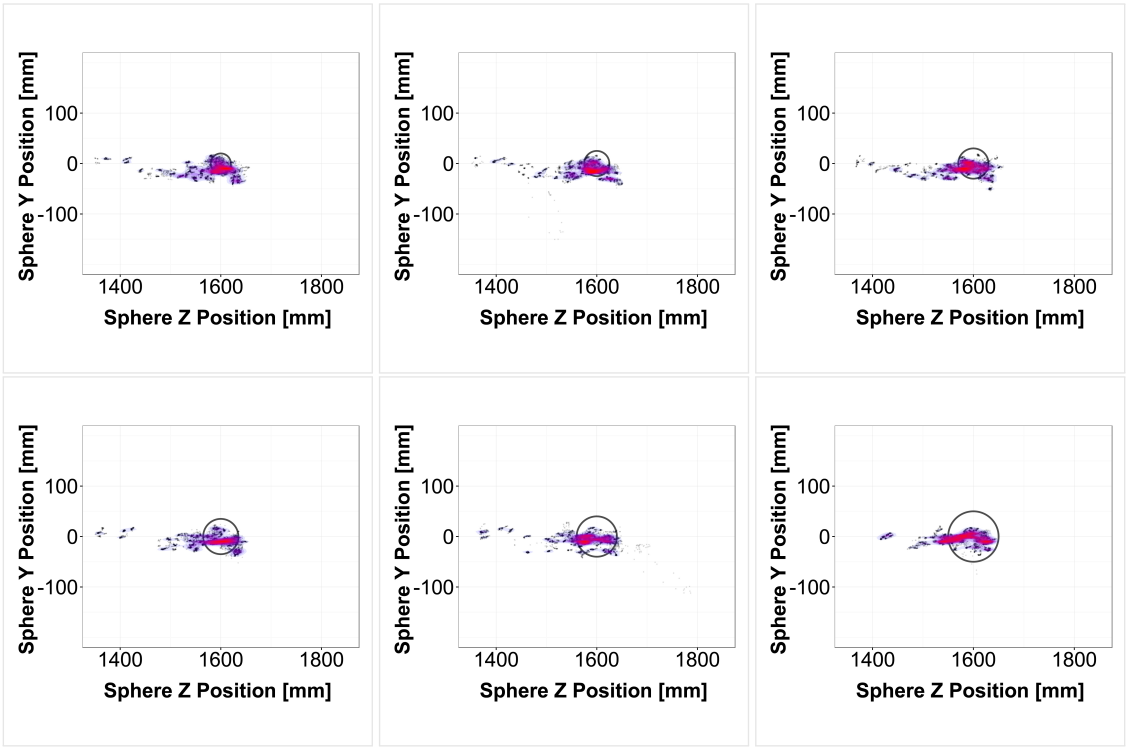
Out of all three axis, $GDisp_{\text{z}}$ presented the highest displacement and variation (see Table 5.3 [page 70]). This was expected as only visual feedback was used through utilising a single monitor, thus users were least confident in their gmp placement along the z axis in comparison to the x and y axis (i.e. $GDisp_{\text{x}}$ and $GDisp_{\text{y}}$).

Negative $GDisp_{\text{z}}$ was found in both objects across all sizes (see Table 5.3 [page 70]), this indicates that majority of participants underestimated the z position of the omp by placing their gmp closer to the sensor in front of the omp for all sizes, as shown in Figures 5.4a [page 75] and 5.4b [page 75]. In this thesis, the terms underestimation and overestimation are opposite to those of depth perception, hence in this study, depth refers to the distance from the feedback monitor and not the user as in depth perception studies.

Overestimation was also present, but not as frequent as underestimation. The position of gmp in the z axis was comparable across all sizes for both objects, mean $GDisp_{\text{z}}$ ranged from 34.34mm to 42.13mm (SD = 61.67) in all cube sizes, and from 27.04mm to 31.86mm (SD = 60.51mm) in all sphere sizes. However, high variation in $GDisp_{\text{z}}$, as shown by the high SD values in Table 5.3



(a) *gmp* placement along the z axis when grasping cubes



(b) *gmp* placement along the z axis when grasping spheres

Figure 5.4: *gmp* placement in the z axis of all participants in six virtual object sizes (in the order 40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 5.4a: black squares indicate cube sizes. 5.4b: black circles indicate sphere sizes. Density heat maps indicate *gmp* placement across participants (red indicates higher density)

[page 70], for both objects was present.
In this experiment grasp variations in terms of orientation and type was present in between participants, indicating that participants adapted their medium wrap grasp that they were instructed to use into different grasp types with changes in object size.

5.3.3.7 Results - Object Type

Statistically significant differences between different object types in different sizes were found in GAp ($\chi^2(1) = 2028$, $p < 0.01$), completion time ($\chi^2(1) = 2926$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 42730$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 50448$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 364$, $p < 0.01$).

5.3.3.8 Analysis - Object Type

In this section, findings for different object types (cubes and spheres) are reported for all object sizes under test, and not for each individual size in this experiment to avoid repetition with results previously reported.

In GAp, users showed a lower mean overestimation of object size in grasping spheres than cubes except for the 80mm and 100mm sizes in which participants performed more accurately in matching the size of the cube than the sphere (see Table 5.3 [page 70]). Participants showed a lower mean GAp in grasping spheres (M = 69.25mm, SD = 28.27) than cubes (M = 77.64mm, SD = 26.36). Mean difference between GAp and object size was lower in 4 of the 6 sizes (40mm, 50mm, 60mm and 70mm) for spheres than cubes. This shows that users were more accurate in matching their GAp to object sizes of spheres, however, users generally presented a GAp working range from 60mm to 88mm regardless of object size or type.

In completion time, participants showed a lower mean completion time in grasping spheres (M = 3.48s, SD = 1.55) than cubes (M = 4.25s, SD = 1.93) across all sizes, where users showed a lower mean completion time in in all six sizes under test when grasping spheres (see Table 5.3 [page 70]). Even though similar ranges and variations of completion times were found for both objects with no clear linear relationship between completion time and object size found, this is in compliance with other findings in this experiment that show that users were more accurate and comfortable when interacting with spheres, leading to shorter completion times for spheres.

In $GDisp_x$, participants showed a lower mean $GDisp_x$ in grasping cubes (M = 28.01mm, SD = 14.08) than spheres (M = 31.51mm, SD = 14.68) across all sizes, where users showed a lower mean $GDisp_x$ in 5 of the 6 sizes (40mm, 50mm, 60mm, 70mm and 80mm) under test. However, SD differences within object sizes between cubes and spheres were comparable (see Table 5.3 [page 70]).

In $GDisp_y$, participants showed a lower mean $GDisp_y$ in grasping spheres (M = -9.84mm, SD = 12.51) than cubes (M = -12.37mm, SD = 11.94) across all sizes, where users showed lower mean $GDisp_y$ in all six sizes under test when grasping spheres. Similar to $GDisp_x$ however, SD

differences in $GDisp_y$ within object sizes between cubes and spheres were again comparable (see Table 5.3 [page 70]).

In $GDisp_z$, participants showed a lower mean $GDisp_z$ in grasping spheres (M = 29.87mm, SD = 60.50) than cubes (M = 38.39mm, SD = 61.67) across all sizes, where users showed lower mean $GDisp_z$ in all six sizes under test when grasping spheres. However, variation gmp placement along the z axis was high for both objects across all sizes (see Table 5.3 [page 70]), thus showing that regardless of object type, users still find it challenging to accurately locate virtual objects in the current setting.

Hypothesis - Revisited

$H_{1.1}$: changes in object size do not have an effect on a) grasp aperture and b) grasp displacement: **Rejected** as GAp is significantly affected by changes in size and object type, within the bounds of the range found (65.70mm to 88.80mm), and $GDisp$ in all three axis ($GDisp_x$, $GDisp_y$ and $GDisp_z$) is also significantly affected by changes in size and object type.

5.4 Experiment 2 - Object Position

5.4.1 Experiment Design

Experiment 2 used a 2 × 9 × 9 × 9 within-subjects design, with two primary conditions: object position and object type. All 15 participants took part in both conditions. Every permutation of position for both object types was randomly presented to participants to exclude potential learning effects. In total, each participant completed 27 (positions) × 2 (objects) = 54 trials and 810 grasps (54 trials × 15 participants). Each static grasp of every participant was recorded for 5 seconds (75 frames), this lead to collecting 60750 raw data points (810 grasps × 75 frames).

5.4.2 Procedure

For this second experiment, participants were instructed to try and accurately find and then accurately match their grasp aperture to the size and position of the virtual object shown to them on the feedback monitor in the shortest time possible. Users were instructed to finish the task in the shortest time possible to test if the test conditions (i.e. virtual object position and type) have any impact on grasp accuracy.
The centre position used in experiment 1 was changed in x, y and z axes. Objects were positioned in 27 different positions in all three axes, covering a range of 400mm away from participants (see Figure 4.7a[page 55]).
Before the experiment, an object (cube or sphere) appeared to participants on the feedback monitor, each object had 27 different positions. The object sizes chosen for this experiment were the two sizes that had the lowest mean difference between GAp and object size in experiment 1 (80mm for cubes and 70mm for spheres) and were unchanged throughout the exper-

iment. A countdown of 5 seconds followed by an auditory cue was used as an indicator for participants to start grasping the object.

5.4.3 Results

Center object position that was used in experiment 1 was changed in this experiment across the x, y and z axis. In order to provide a valid and direct comparison between the two experiments in this study, this section will only report on and analyse results of the z plane that was used in experiment 1 (1600mm), and changes in position were compared to the centre position to asses the influence of position changes on GAp and $GDisp$. The influence of changes in the z plane on GAp and $GDisp$ is analysed in the form of set comparisons and not individual positions. Full comparisons of all positions across all z planes are reported in Table 5.4 [page 79].

5.4.3.1 Results - Grasp Aperture (GAp)

1400m Z plane

Statistically significant differences in GAp between different cube ($\chi^2(8) = 632$, $p < 0.01$) and sphere ($\chi^2(8) = 1533$, $p < 0.01$) positions were found. Significant adjusted post-hoc results are reported in Table 5.4 [page 79] (see ● for cubes, ○ for spheres).

1600m Z plane

Statistically significant differences in GAp between different cube ($\chi^2(8) = 559$, $p < 0.01$) and sphere ($\chi^2(8) = 2144$, $p < 0.01$) positions were found.

1800m Z plane

Statistically significant differences in GAp between different cube ($\chi^2(8) = 1397$, $p < 0.01$) and sphere ($\chi^2(8) = 1785$, $p < 0.01$) positions were found.

5.4.3.2 Analysis - Grasp Aperture (GAp)

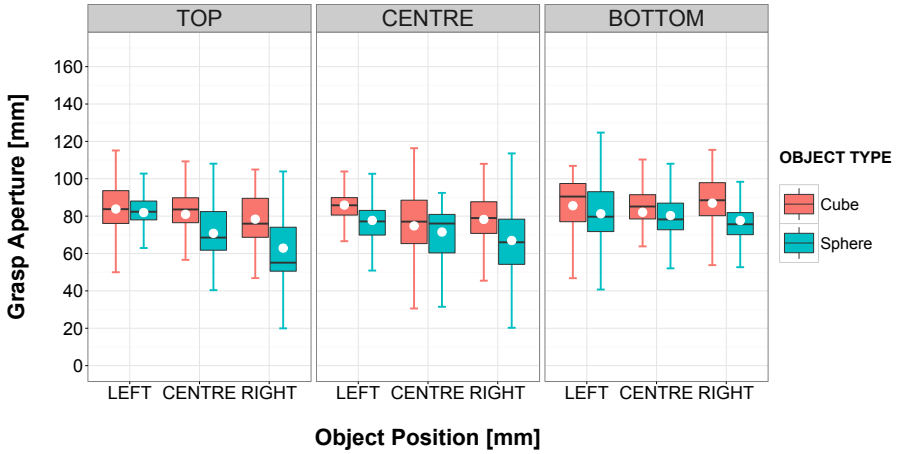
Similar to experiment 1, users overestimated object size in the majority of positions in this experiment, this was consistent across all z planes where mean overestimation occurred in 33 out of the 54 trials in this experiment (see Table 5.5 [page 81]). Users also showed similar ranges of mean GAp across all z planes for cubes (from 64.46 $\pm$ 18.39 to 90.60 $\pm$ 17.35) and spheres (from 62.27 $\pm$ 15.33 to 90.06 $\pm$ 16.27), despite the constant sizes of the two objects in this experiment. This shows that users were again grasping within a specific range (60mm to 90mm) regardless of object size. This high variation in size estimation using GAp can be attributed to the lack of tactile feedback in freehand grasping, and to the additional task of locating virtual objects in this experiment, potentially impacted grasping performance in terms of matching GAp to object size due to the additional cognitive load. Participants changed their GAp in the majority of position changes of objects, however similarity of grasps between participants was also found in some positions (e.g. Centre, Centre Right and Top Right in the 1600mm z plane as shown in Table 5.4 [page 79]) that potentially show a working range or region which is preferable by participants.

Table 5.4: Significant Post-hoc Pairwise Comparisons - Experiment 2. Symbols are represented in constant order (● ○ ▲ △ ◆ ◇ ■ □) and represent significance in a post-hoc Dunn Test with Bonferroni correction using an α level of 0.01 for the following: ● GAp - Cube, ○ GAp - Sphere, ▲ $GDisp_{\mathrm{x}}$ - Cube, △ $GDisp_{\mathrm{x}}$ - Sphere, ◆ $GDisp_{\mathrm{y}}$ - Cube, ◇ $GDisp_{\mathrm{y}}$ - Sphere, ■ $GDisp_{\mathrm{z}}$ - Cube, □ $GDisp_{\mathrm{z}}$ - Sphere. No symbols indicate statistical similarity

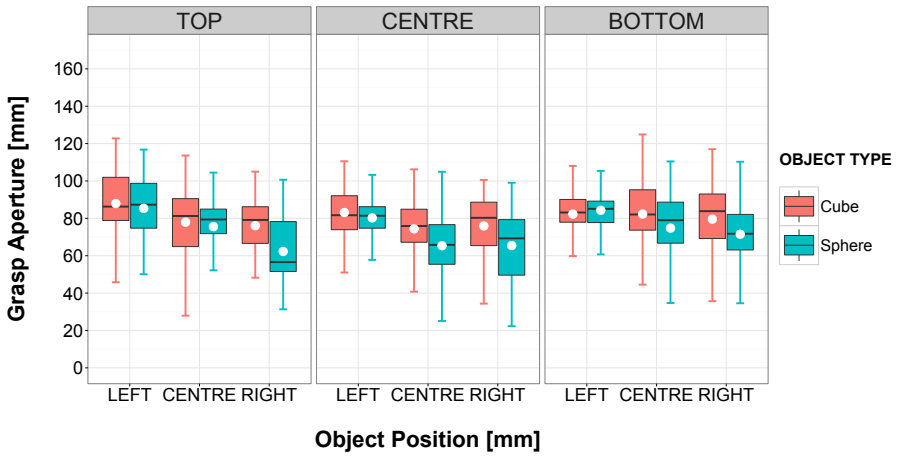
Z Plane 1 - 1400mm									
Positions	Top Left	Top Centre	Top Right	Centre Left	Centre	Centre Right	Bottom Left	Bottom Centre	Bottom Right
Top Left		○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ▲ △ ◆ ◇ ■ □	○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Top Centre			● ○ ▲ △ ◇ ■ □	● ○ △ ◆ ◇ □	● ▲ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Top Right				● ○ ▲ △ ◆ ◇	○ ▲ △ ◆ ◇	○ ▲ △ ◆ ◇ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ □	● ○ ▲ △ ◆ ◇ ■ □
Centre Left					● ○ △ ◆ ◇	● ○ ▲ △ ◇ □	○ ◆ ◇ ■ □	○ ▲ △ ◆ ◇ □	▲ △ ◆ ◇ ■ □
Centre						○ ▲ △ ◆ ◇ □	● ○ △ ◆ ◇ □	● ○ ◆ ◇ □	● ○ ▲ △ ◆ ◇ ■ □
Centre Right							● ○ ▲ △ ◆ ◇ □	● ○ ▲ △ ◆ ◇ □	● ○ ◆ ◇ ■ □
Bottom Left								● ▲ △ ◆	○ ▲ △ ◆ ■ □
Bottom Centre									● ○ ▲ △ ■
Bottom Right									

Z Plane 2 - 1600mm									
Positions	Top Left	Top Centre	Top Right	Centre Left	Centre	Centre Right	Bottom Left	Bottom Centre	Bottom Right
Top Left		● ○ ▲ △ ◆ ◇ ■	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Top Centre			● ○ ▲ △ ◆ ■ □	● ○ △ ◆ ◇ ■ □	● ○ △ ◆ ◇ □	○ ▲ △ ◆ ◇ ■ □	● ○ △ ◆ ◇ ■ □	● △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Top Right				● ○ ▲ △ ◆ ◇	▲ △ ◆ ◇ ■	○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Centre Left					● ○ △ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	○ ◆ ◇ ■ □	○ ▲ △ ◆ ◇ ■ □	○ ▲ △ ◆ ◇ ■ □
Centre						▲ △ ◆ ◇ ■ □	● ○ △ ◆ ◇ ■ □	● ○ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Centre Right							● ○ ▲ △ ◆ ◇ □	● ○ ▲ △ ◆ ◇ □	● ○ ▲ ◆ ◇ ■ □
Bottom Left								○ ▲ △ ◇ ■	○ ▲ △ ◆ ◇ ■ □
Bottom Centre									○ ▲ △ ◆ □
Bottom Right									

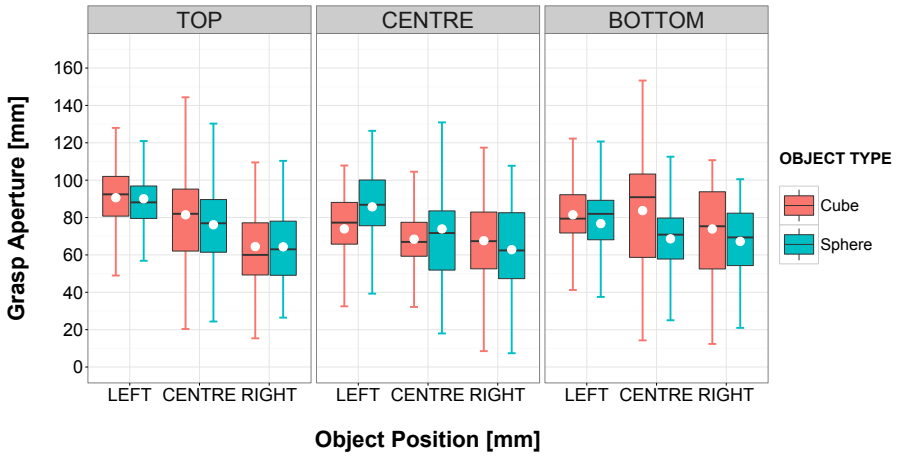
Z Plane 3 - 1800mm									
Positions	Top Left	Top Centre	Top Right	Centre Left	Centre	Centre Right	Bottom Left	Bottom Centre	Bottom Right
Top Left		● ○ ▲ △ ◇ ■ □	● ○ ▲ △ ◇ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Top Centre			● ○ ▲ △ ◇ ■	● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ □	▲ △ ◆ ◇ ■ □	● ○ ▲ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Top Right				● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ □	● ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■	● ○ ▲ △ ◆ ◇ ■	● ○ ▲ △ ◆ ◇ ■ □
Centre Left					● ○ ▲ △	● ○ ▲ △ ◇ □	● ○ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	○ ▲ △ ◆ ◇ ■ □
Centre						○ ▲ △ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □
Centre Right							● ○ ▲ △ ◆ ◇ ■ □	● ○ ▲ △ ◆ ◇ ■ □	● ▲ △ ◆ ◇ ■
Bottom Left								○ ▲ △ ◆	● ○ ▲ △ ◆ ■ □
Bottom Centre									● ▲ △ ◆ ■ □
Bottom Right									



(a) 1400mm z plane



(b) 1600mm z plane



(c) 1800mm z plane

Figure 5.5: GAp for different object positions in the three z planes in this experiment. 5.6c: 1400mm z plane. 5.6b: 1600mm z plane. 5.6a: 1800mm z plane. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

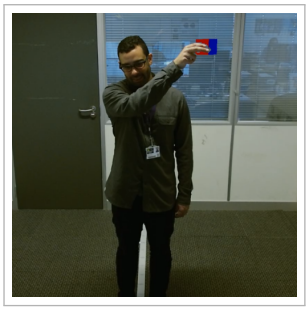
Table 5.5: Descriptive Statistics of GAp for different positions of cubes and spheres (Mean ± SD). For statistical significance (p < 0.01) between individual positions in each z plane, see Table 5.4 [page 79]

1400mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cube (Constant Size - 80mm)	Top	83.82 ± 15.30	80.97 ± 13.84	78.40 ± 13.15
	Centre	85.92 ± 10.98	74.77 ± 19.50	78.24 ± 13.95
	Bottom	85.58 ± 13.90	82.06 ± 20.11	86.95 ± 14.88
Spheres (Constant Size - 70mm)	Top	81.92 ± 9.80	70.80 ± 14.69	62.89 ± 17.82
	Centre	77.71 ± 13.15	71.51 ± 11.95	66.96 ± 19.53
	Bottom	81.27 ± 14.58	80.30 ± 10.82	77.64 ± 12.96
1600mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	87.91 ± 17.41	77.96 ± 17.00	76.04 ± 14.74
	Centre	83.13 ± 12.39	74.36 ± 16.84	75.92 ± 15.10
	Bottom	82.28 ± 13.79	82.31 ± 15.33	79.56 ± 17.45
Spheres (Constant Size - 70mm)	Top	85.36 ± 15.59	75.54 ± 16.69	62.27 ± 15.33
	Centre	80.25 ± 10.23	65.34 ± 15.76	65.44 ± 17.66
	Bottom	84.33 ± 14.58	74.77 ± 18.68	71.50 ± 16.72
1800mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	90.60 ± 17.35	81.46 ± 26.56	64.46 ± 18.39
	Centre	73.93 ± 19.51	68.41 ± 19.09	67.62 ± 20.60
	Bottom	81.46 ± 20.46	83.71 ± 26.17	73.88 ± 22.48
Spheres (Constant Size - 70mm)	Top	90.06 ± 16.27	76.13 ± 20.92	64.33 ± 24.46
	Centre	85.72 ± 20.27	73.91 ± 31.14	62.81 ± 20.80
	Bottom	76.75 ± 20.32	68.62 ± 21.99	67.23 ± 16.18

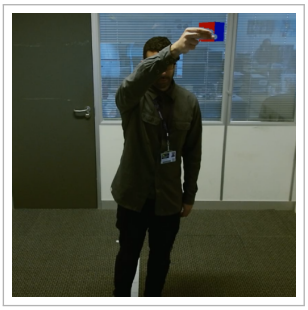
Statistically significant differences in GAp between different z planes were found for cubes (χ^2 (2) = 458, p < 0.01), but not spheres (χ^2 (2) = 3.63, p > 0.01). This shows that participants altered their GAp as position of objects changed in the z axis when grasping cubes, but showed comparable GAp across all z planes when grasping spheres. Users presented the highest accuracy in terms of matching their GAp to object sizes in the 1600mm z plane for cubes (mean underestimation of -0.06mm ± 16.17) and spheres (mean overestimation of 3.87mm ± 17.73) across all positions, this preference for the middle z plane is potentially due to the convenience of its spatial position in relation to the biomechanical reach of users, as interaction in this plane does not require as much flexion and extension of the forearm in comparison to the 1400mm (furthest from users) and 1800mm (closest to users) that represent the extremities of the mean reach of users, and can be physically more demanding in terms of reaching for virtual objects and the subsequent fine grasp adjustments (see Figure 5.6 [page 82]).

As shown in Figures 5.8f [page 83] and 5.5 [page 80], users showed the highest accuracy in

(a) 1800mm z Plane

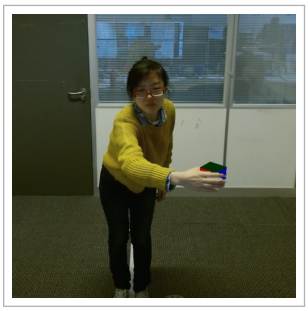


(b) 1600mm z Plane



(c) 1400mm z Plane

Figure 5.6: Users showed higher accuracy in terms of matching GAp to objects size in the middle 1600mm z plane (5.6b) due to the spatial convenience of this position for users



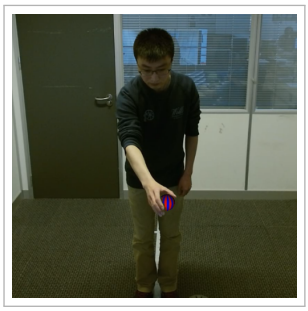


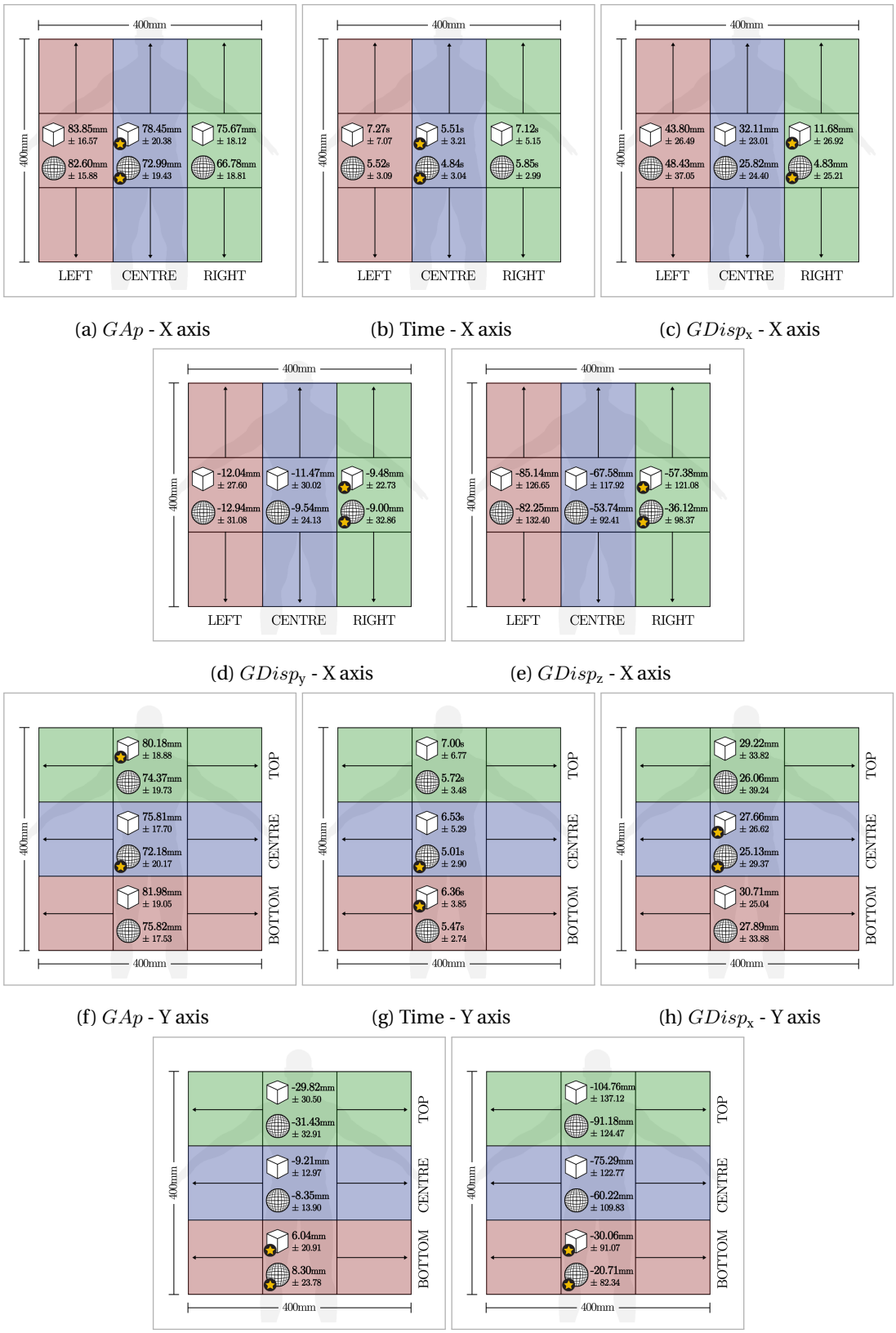
Figure 5.7: Bottom positions were problematic for users in size estimation, especially in the furthest z plane from users where leaning forward was required to correctly grasp the object

matching GAp to object size across all z planes in the Top positions alongside the y axis (Top Left, Top Centre and Top Right) for cubes with a mean overestimation of 0.18mm $\pm$ 18.88, and in Centre positions (Centre Left, Centre and Centre Right) for spheres with a mean overestimation of 2.18mm $\pm$ 20.17. This accuracy in Top and Centre positions is potentially due to their easily accessible positions in relation to the height of users, unlike Bottom positions that may lead some users to bend their backs or lean forward to be able to grasp and accurately match size of objects using their GAp (especially in the furthest z plane from users), that can hinder their estimation due to the inconvenience of the position of the interaction region (see Figure 5.7 [page 82]).

Along the x axis (see Figures 5.8a [page 83]), users presented the highest accuracy in matching GAp to object size across all z planes in the Centre positions (Centre Left, Centre, Centre Right) for cubes (mean underestimation of -1.55mm $\pm$ 20.38) and spheres (mean overestimation of 2.99mm $\pm$ 19.43). As shown in Figure 5.5 [page 80], users were generally more accurate in matching GAp to object sizes in Right and Centre positions, this was expected as all users in this study were right-handed.

5.4.3.3 Results - Completion Time

1400mm Z plane



(a) GAp - X axis (b) Time - X axis (c) $GDisp_{x}$ - X axis

(d) $GDisp_{y}$ - X axis (e) $GDisp_{z}$ - X axis

(f) GAp - Y axis (g) Time - Y axis (h) $GDisp_{x}$ - Y axis

(i) $GDisp_{y}$ - Y axis (j) $GDisp_{z}$ - Y axis

Figure 5.8: Optimal interaction regions for users across all z planes in terms of all the measurement used in this work to assess grasp accuracy. X axis refers to Left, Centre and Right positions. Y axis refers to Top, Centre and Bottom positions. Values presented are Means ± SD of each corresponding measurement. Most accurate/quickest region is marked with a star

Statistically significant differences in completion time between different cube ($\chi^2(8) = 673$, $p < 0.01$) and sphere ($\chi^2(8) = 755$, $p < 0.01$) positions were found.

1600mm Z plane

Statistically significant differences in completion time between different cube ($\chi^2(8) = 380$, $p < 0.01$) and sphere ($\chi^2(8) = 739$, $p < 0.01$) positions were found.

1800mm Z plane

Statistically significant differences in completion time between different cube ($\chi^2(8) = 257$, $p < 0.01$) and sphere ($\chi^2(8) = 439$, $p < 0.01$) positions were found.

5.4.3.4 Analysis - Completion Time

Table 5.6: Descriptive Statistics of Task Completion Time (Mean ± SD)

1400mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cube (Constant Size - 80mm)	Top	7.20 ± 5.67	5.46 ± 2.25	6.60 ± 3.65
	Centre	5.93 ± 3.34	5.27 ± 2.18	9.20 ± 6.15
	Bottom	6.67 ± 3.61	5.66 ± 2.09	9.20 ± 5.78
Spheres (Constant Size - 70mm)	Top	6.40 ± 3.24	4.46 ± 2.03	7.53 ± 4.15
	Centre	5.20 ± 2.54	5.60 ± 4.50	5.73 ± 2.44
	Bottom	5.73 ± 2.98	5.20 ± 2.29	6.87 ± 3.58
1600mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	6.73 ± 5.79	5.93 ± 4.11	8.46 ± 8.46
	Centre	6.93 ± 7.26	5.33 ± 3.34	6.00 ± 4.73
	Bottom	6.00 ± 3.88	5.00 ± 3.74	6.27 ± 2.54
Spheres (Constant Size - 70mm)	Top	5.80 ± 3.37	5.13 ± 3.33	4.60 ± 1.36
	Centre	4.53 ± 2.68	4.46 ± 1.96	5.00 ± 2.39
	Bottom	5.80 ± 2.20	4.53 ± 2.06	6.40 ± 2.89
1800mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	11.40 ± 13.80	5.07 ± 2.77	6.13 ± 3.44
	Centre	7.93 ± 8.12	6.13 ± 4.17	6.07 ± 3.91
	Bottom	6.60 ± 3.85	5.73 ± 3.22	6.13 ± 3.33
Spheres (Constant Size - 70mm)	Top	6.00 ± 4.17	5.93 ± 4.78	5.67 ± 2.41
	Centre	4.73 ± 2.86	4.06 ± 2.05	5.80 ± 3.29
	Bottom	5.47 ± 2.90	4.20 ± 2.04	5.06 ± 2.24

Mean completion time time ranged across all z planes from 5.00s ± 3.74 to 11.40s ± 13.80 for cubes, and from 4.06s ± 2.05 to 7.53s ± 4.15 for spheres (see Table 5.6 [page 84]). Variation in completion times between users can be attributed to individual differences in perceptual tasks (e.g. users that are experienced in AR interaction) (see Section 3.3.2 [page 32]).

Users showed shortest mean completion times in the 1600mm z plane when grasping both cubes (6.30s ± 5.29) and spheres (5.14s ± 2.63) (see Table 5.6 [page 84]). This preference for the 1600mm z plane can again be attributed to the convenience of its position in relation to the biomechanical reach of users, as this plane is not as physically demanding as the furthest and closest z planes.

As shown in Figure 5.8g [page 83], users also showed shortest mean completion times across all planes in the Bottom positions alongside the y axis across all z planes for cubes (mean completion time of 6.36s ± 3.85), and in Centre positions for spheres (mean completion time of 5.01s ± 2.90). Alongside the x axis (see Figure 5.8b [page 83]), users showed shortest mean completion times across all z planes in the Centre positions for cubes (mean completion time of 5.51s ± 3.21) and spheres (mean completion time of 4.84s ± 3.04). Central interaction regions across the x axis are usually preferred by users as they are easily accessible using the right dominant hand and reach to grasp movements.

Even though statistically significant differences in completion time between different cube positions and sphere positions were found, and no trends between completion time and object size were found.

5.4.3.5 Results - Grasp Displacement ($GDisp$)

Analysis of results in this section is between z planes as full sets, full significant adjusted post-hoc results for individual positions are reported in Table 5.4 [page 79].

1400mm Z plane

Statistically significant differences between different object positions were found in $GDisp_{\text{x}}$ (cube ($\chi^2(8) = 1333$, $p < 0.01$) and sphere ($\chi^2(8) = 2106$, $p < 0.01$)), $GDisp_{\text{y}}$ (cube ($\chi^2(8) = 3680$, $p < 0.01$) and sphere ($\chi^2(8) = 3591$, $p < 0.01$)) and $GDisp_{\text{z}}$ (cube ($\chi^2(8) = 1125$, $p < 0.01$) and sphere ($\chi^2(8) = 1805$, $p < 0.01$)).

1600mm Z plane

Statistically significant differences between different object positions were found in $GDisp_{\text{x}}$ (cube ($\chi^2(8) = 1954$, $p < 0.01$) and sphere ($\chi^2(8) = 3251$, $p < 0.01$)), $GDisp_{\text{y}}$ (cube ($\chi^2(8) = 3873$, $p < 0.01$) and sphere ($\chi^2(8) = 4174$, $p < 0.01$)) and $GDisp_{\text{z}}$ (cube ($\chi^2(8) = 1218$, $p < 0.01$) and sphere ($\chi^2(8) = 1455$, $p < 0.01$)).

1800mm Z plane

Statistically significant differences between different object positions were found in $GDisp_{\text{x}}$ (cube ($\chi^2(8) = 3694$, $p < 0.01$) and sphere ($\chi^2(8) = 3933$, $p < 0.01$)), $GDisp_{\text{y}}$ (cube positions ($\chi^2(8) = 4019$, $p < 0.01$) and sphere positions ($\chi^2(8) = 4074$, $p < 0.01$)) and $GDisp_{\text{z}}$ (cube ($\chi^2(8) = 1335$, $p < 0.01$) and sphere ($\chi^2(8) = 1383$, $p < 0.01$)).

5.4.3.6 Analysis - Grasp Displacement ($GDisp$)

$GDisp_{\text{x}}$

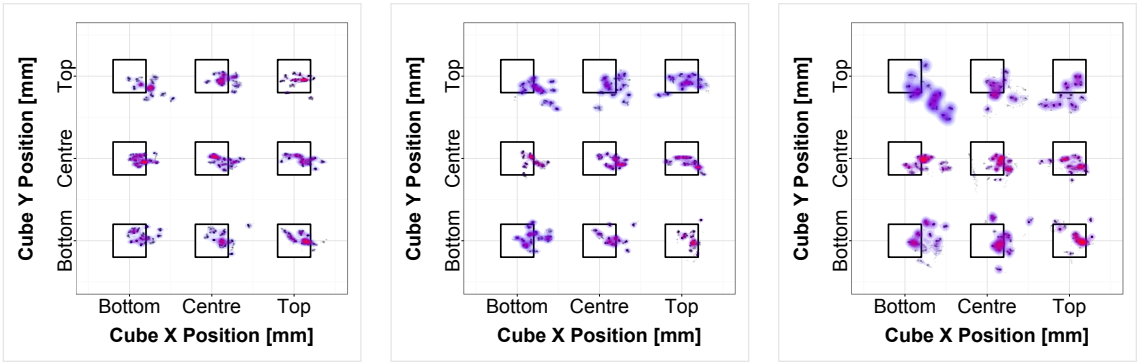
Table 5.7: Descriptive Statistics of $GDisp_x$ (Mean ± SD)

1400mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cube (Constant Size - 80mm)	Top	45.82 ± 25.96	33.00 ± 20.56	10.53 ± 22.32
	Centre	30.68 ± 17.96	28.28 ± 22.35	20.11 ± 26.80
	Bottom	31.69 ± 22.53	27.28 ± 24.10	17.44 ± 21.94
Spheres (Constant Size - 70mm)	Top	45.30 ± 27.97	23.56 ± 20.07	1.03 ± 23.67
	Centre	28.91 ± 17.67	21.64 ± 17.28	13.07 ± 27.16
	Bottom	42.71 ± 45.76	20.22 ± 24.01	12.99 ± 18.57
1600mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	52.28 ± 26.93	34.77 ± 23.36	5.94 ± 30.31
	Centre	37.44 ± 18.57	34.81 ± 19.01	13.49 ± 25.95
	Bottom	38.82 ± 22.70	32.75 ± 25.12	23.18 ± 14.83
Spheres (Constant Size - 70mm)	Top	47.93 ± 33.40	21.88 ± 17.66	-2.96 ± 24.53
	Centre	40.44 ± 25.24	25.64 ± 17.02	7.32 ± 25.95
	Bottom	38.71 ± 27.93	31.17 ± 26.45	8.75 ± 14.61
1800mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	58.82 ± 31.81	27.29 ± 29.03	-5.47 ± 28.86
	Centre	48.19 ± 29.52	33.71 ± 16.12	2.24 ± 28.51
	Bottom	50.45 ± 24.43	37.12 ± 22.81	17.63 ± 26.03
Spheres (Constant Size - 70mm)	Top	78.03 ± 41.89	27.33 ± 32.94	-7.56 ± 32.83
	Centre	57.21 ± 43.80	27.78 ± 19.05	4.13 ± 18.35
	Bottom	56.63 ± 36.94	33.16 ± 34.30	6.69 ± 28.07

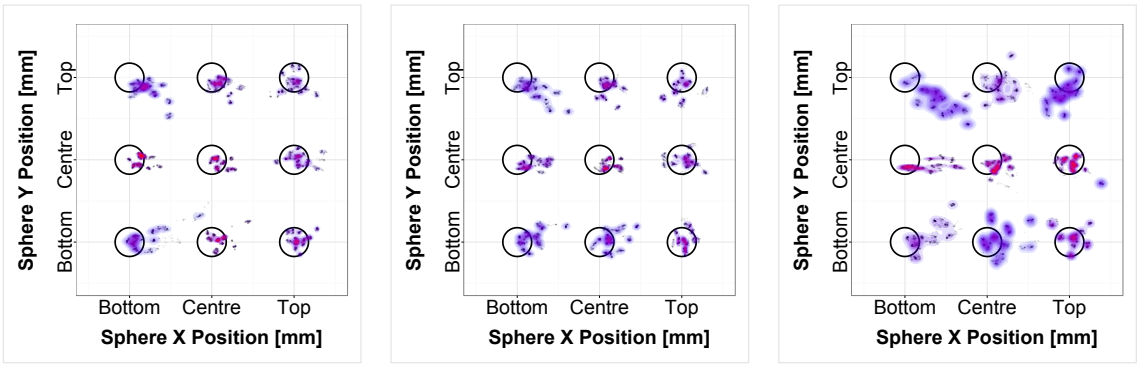
$GDisp_x$ showed the second highest displacement in this study for both objects. Users showed a mean $GDisp_x$ of 29.20mm ± 28.78 in grasping cubes, and a mean of 26.35mm ± 34.42 for spheres across all positions and z planes.

Users showed high variation in gmp placement along the x axis as statistical differences were found in $GDisp_x$ between the majority of positions for both objects (see Table 5.4 [page 79]). However participants showed statistically similar $GDisp_x$ to the center position used in the experiment 1 in some positions (e.g. top centre, centre left, bottom left, bottom centre and bottom right positions of the cube in the 1600mm z plane - see Table 5.4 [page 79]), this potentially shows preferable interaction regions by users in freehand grasping.

Mean $GDisp_x$ ranged from -59.12mm to 100.12mm (SD = 26.90) for cubes, and from -53.54mm to 127.40mm (SD = 29.19) for spheres as shown by clusters in Figures 5.9a [page 87] and 5.9b [page 87]. This range for both objects shows that even though participants showed statistical similarities in gmp placement along the x axis in some positions, $GDisp_x$ shows high variation between users. Users showed statistically significant differences in $GDisp_x$ between different z planes for cubes ($\chi^2(2) = 114$, $p < 0.01$) and spheres ($\chi^2(2) = 162$, $p < 0.01$), and presented the



(a) gmp placement along the x and y axes when grasping cubes in 27 positions in 3 z planes



(b) gmp placement along the x and y axes when grasping spheres in 27 positions in 3 z planes

Figure 5.9: gmp placement in the x and y axes of all participants in 27 positions in 3 z planes (in the order: 1400mm z plane - 1600mm z plane - 1800mm z plane). 5.9a: black squares indicate cube sizes. 5.9b: black circles indicate sphere sizes. Density heat maps indicate gmp placement across participants (red indicates higher density)

lowest mean $GDisp_{\text{x}}$ in the 1400mm z plane for cubes (27.20mm ± 24.81) and spheres (23.27mm ± 29.28) across all positions (see Table 5.7 [page 86]). This preference for the furthest z plane from users can be attributed to the position of this particular z plane that is at the extremity of the mean arm reach of users, this extreme position of the plane naturally limited any errors in gmp placement along the x axis due to the limited possible movements using the hand.
As shown in Figure 5.8h [page 83], lowest mean $GDisp_{\text{x}}$ was shown by users in the Centre positions alongside the y axis across all z planes for cubes (27.66 ± 26.62) and spheres (25.13 ± 29.37). Alongside the x axis (see Figure 5.8c [page 83]), users showed the lowest $GDisp_{\text{x}}$ in Right positions across all z planes for cubes (11.68 ± 26.92) and spheres (4.83 ± 25.21), this preference for right positions is potentially attributed to the right handedness of all users in this study, as users showed highest mean $GDisp_{\text{x}}$ in Left positions for both objects (see Table 5.7 [page 86]).

***GDisp*$_{\text{y}}$**

$GDisp_{\text{y}}$ showed the lowest mean grasp displacement in this study across all positions and z planes for cubes (-11.00 ± 26.98) and spheres (-10.49 ± 29.65) for both objects. Similar to findings in experiment 1, negative $GDisp_{\text{y}}$ was found in the majority of positions for both objects, where users placed their gmp to a point that is lower than the omp in 35 out of the 54 trials in

Table 5.8: Descriptive Statistics of $GDisp_{\mathrm{y}}$ (Mean ± SD)

1400mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cube (Constant Size - 80mm)	Top	-27.76 ± 16.51	-9.91 ± 12.09	-12.06 ± 15.73
	Centre	-3.65 ± 9.07	-6.45 ± 10.51	-3.63 ± 9.67
	Bottom	11.59 ± 13.15	4.56 ± 17.62	4.76 ± 11.45
Spheres (Constant Size - 70mm)	Top	-31.00 ± 21.58	-19.24 ± 16.07	-14.40 ± 21.35
	Centre	-2.54 ± 12.15	-9.26 ± 11.96	1.25 ± 13.35
	Bottom	8.53 ± 23.16	6.30 ± 17.19	6.08 ± 16.25
1600mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	-34.62 ± 20.65	-27.75 ± 22.98	-16.69 ± 15.20
	Centre	-7.64 ± 13.67	-7.60 ± 10.99	-10.60 ± 11.50
	Bottom	7.72 ± 17.47	6.60 ± 17.16	2.35 ± 20.22
Spheres (Constant Size - 70mm)	Top	-35.15 ± 23.21	-26.81 ± 19.04	-26.98 ± 21.39
	Centre	-9.82 ± 10.74	-13.24 ± 10.97	-2.96 ± 15.14
	Bottom	10.29 ± 20.18	5.89 ± 21.00	5.84 ± 19.84
1800mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	-47.63 ± 34.09	-47.82 ± 53.81	-44.18 ± 26.56
	Centre	-12.69 ± 12.53	-15.61 ± 16.18	-14.98 ± 14.37
	Bottom	6.31 ± 26.34	0.73 ± 32.73	9.75 ± 21.34
Spheres (Constant Size - 70mm)	Top	-54.92 ± 30.26	-23.32 ± 29.25	-51.04 ± 62.83
	Centre	-15.08 ± 10.01	-15.62 ± 15.78	-7.89 ± 13.27
	Bottom	13.19 ± 32.93	9.43 ± 32.22	9.11 ± 24.11

this experiment (see Table 5.8 [page 88]), this technique is implemented by users to avoid fully occluding the virtual object presented using their grasp, and to fully, or partially, visualise the object during grasping to verify the validity or accuracy of their grasp on the feedback monitor. Similar to $GDisp_{\mathrm{x}}$, users showed high variation in gmp placement along the y axis as statistical differences were found in $GDisp_{\mathrm{y}}$ between the majority of positions for both objects (see Table 5.4 [page 79]). Participants also showed statistically similar $GDisp_{\mathrm{y}}$ to the center position used in the experiment 1 in some positions (e.g. centre and centre left positions of the cube in the 1600mm z plane - see Table 5.4 [page 79]), this again potentially shows preferable interaction regions by users in freehand grasping.

As shown by clusters in Figures 5.9a [page 87] and 5.9b [page 87], mean $GDisp_{\mathrm{y}}$ ranged from -83.48mm to 52.02mm (SD = 21.90) for cubes, and from -88.41mm to 59.70mm (SD = 24.13) for spheres. Wide ranges across participants and objects show variability in gmp placement in the y axis, and highlights the impact of individual differences on gmp placement (see Section 3.3.2 [page 32]).

Users showed statistically significant differences in $GDisp_{\mathrm{y}}$ between different z planes for cubes

(χ^2 (2) = 1286, p < 0.01) and sphere (χ^2 (2) = 472, p < 0.01), and presented the lowest mean $GDisp_y$ in the 1400mm z plane for cubes (-4.73mm ± 17.08) and spheres (-6.03mm ± 21.59) across all positions (see Table 5.8 [page 88]). Similar to $GDisp_x$ this preference for the furthest z plane from users is attributed to the position of this particular z plane that is at the extremity of the mean arm reach of users.

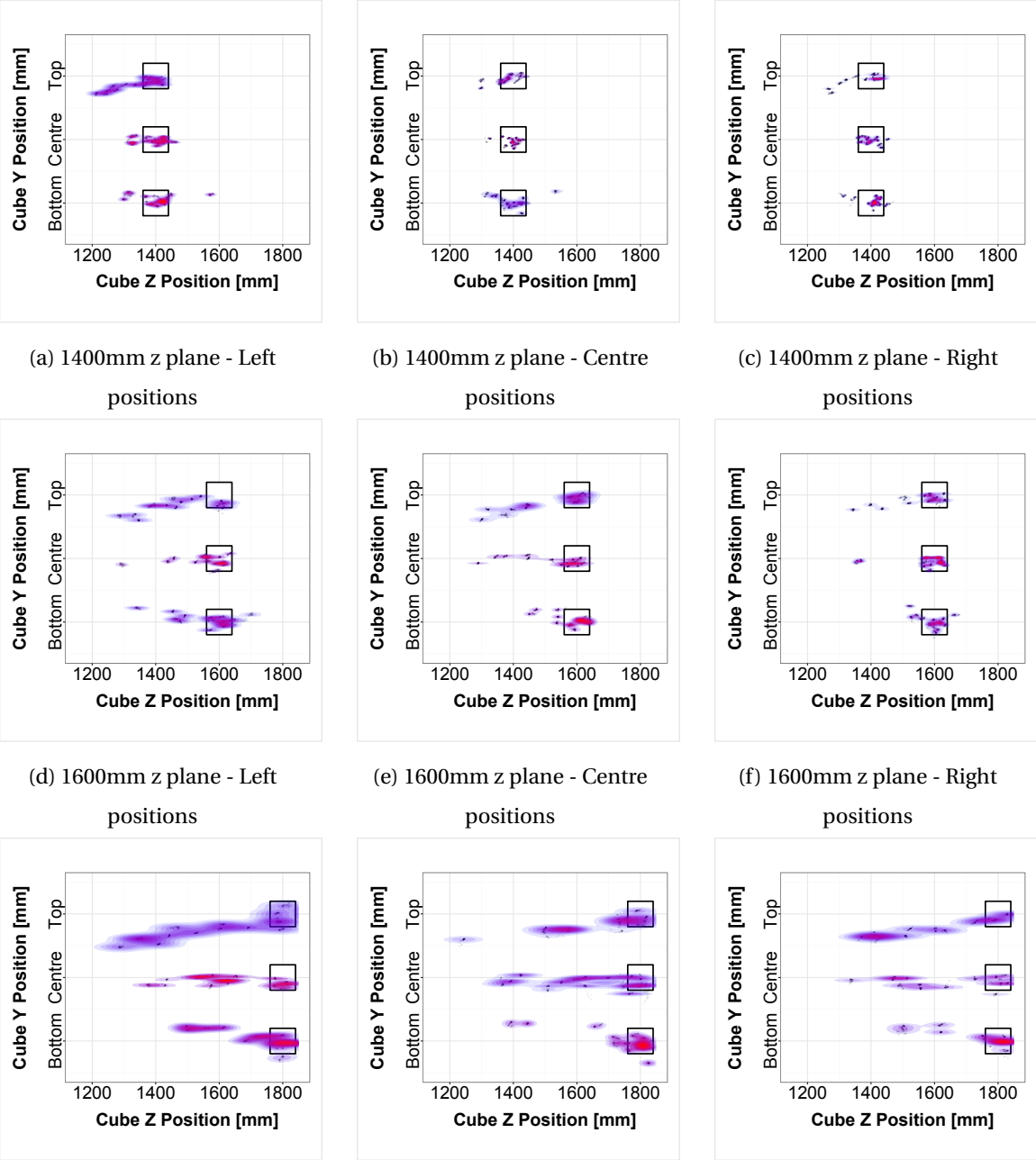
As shown in Figure 5.8i [page 83], lowest mean $GDisp_y$ was shown by users in the Bottom positions alongside the y axis across all z planes for cubes (6.04 ± 20.91) and spheres (8.30 ± 23.78). Bottom positions are potentially preferred by users in gmp placement along the y axis as the arm position during grasping of objects placed in bottom positions is not blocking the users view of the feedback monitor in anyway, unlike Top and Centre positions where the position of the arm can be intersecting the view of users of the feedback monitor thus hindering gmp placement along the y axis. Alongside the x axis (see Figure 5.8d [page 83]), users showed the lowest $GDisp_y$ in Right positions across all z planes for cubes (-9.48 ± 22.73) and spheres (-9.00 ± 32.85) (see Table 5.8 [page 88]), similar to $GDisp_x$ this preference for right positions is attributed to the right handedness of all users in this study.

***GDisp*$_z$**

$GDisp_z$ showed the highest mean grasp displacement in this study across all positions and z planes for cubes (-70.05 ± 122.47) and spheres (-57.37 ± 110.80) (see Table 5.9 [page 92]).

Users showed high variation in gmp placement along the z axis as statistical differences were found in $GDisp_z$ between the majority of positions for both objects (see Table 5.4 [page 79]). Changes in position have noticeably increased the mean $GDisp_z$ in comparison to experiment 1 for cubes (from -105.87mm to 320.06mm (SD = 94.90)) and spheres (from -88.79mm to 323.09mm (SD = 89.06)). Depth estimation in freehand grasping is problematic due to lack of tactile feedback and users were least confident in their gmp placement along the z axis in comparison to the x and y axes, this is evident by the high SD values in depth estimation in different positions (see Table 5.9 [page 92]).

As shown in Figures 5.10 [page 90] and 5.11 [page 91], majority of participants have underestimated the position of objects in the z axis, where depth underestimation occurred in 48 out of the 54 trials in this experiment (see Table 5.9 [page 92]). Meaning that participants consistently placed their gmp in front of the omp regardless of the object type or position. This underestimation has also shown to decrease as objects were further away from participants (1400mm z plane), where users presented the lowest mean $GDisp_z$ in this z plane for cubes (-13.30 ± 52.11) and spheres (-10.61 ± 56.19). It can be argued that participants were more accurate in depth estimation as objects were further away from them, however margin of error in the furthest z plane was limited as the mean arm length of participants was 548mm, this significantly reduced $GDisp_z$, as more depth estimation would be outside the biomechanical arm reach of participants. Users showed statistically significant differences in $GDisp_z$ between different z planes for cubes (χ^2 (2) = 3376, p < 0.01), and spheres (χ^2 (2) = 3104, p < 0.01). This shows that participants altered their gmp placement as the position of objects changed in the z axis.



(a) 1400mm z plane - Left positions

(b) 1400mm z plane - Centre positions

(c) 1400mm z plane - Right positions

(d) 1600mm z plane - Left positions

(e) 1600mm z plane - Centre positions

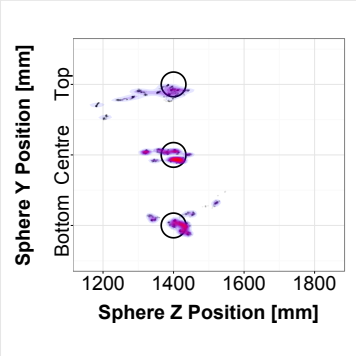
(f) 1600mm z plane - Right positions

(g) 1400mm z plane - Left positions

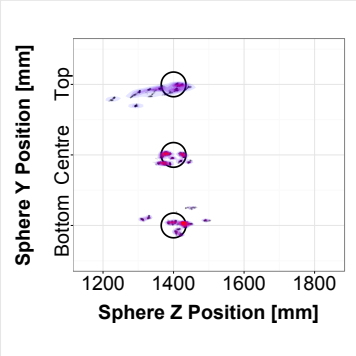
(h) 1400mm z plane - Centre positions

(i) 1400mm z plane - Right positions

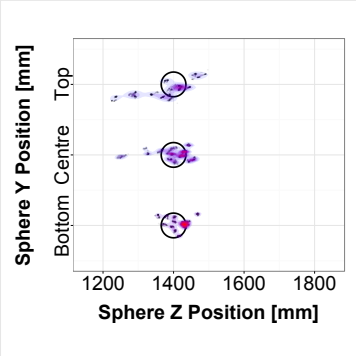
Figure 5.10: gmp placement in the z axis of all participants for gras[ing cubes in 27 positions in 3 z planes (First row: 1400mm z plane. Second row: 1600mm z plane. Third row: 1800mm z plane): squares indicate cube positions (Left, Centre and Right), and density heat maps indicate gmp placement across participants (red indicates higher density)



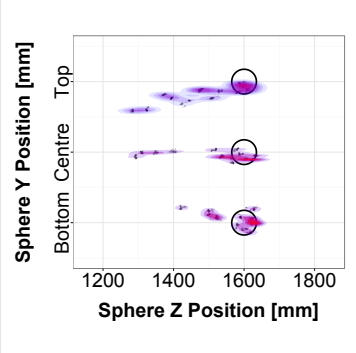
(a) 1400mm z plane - Left positions



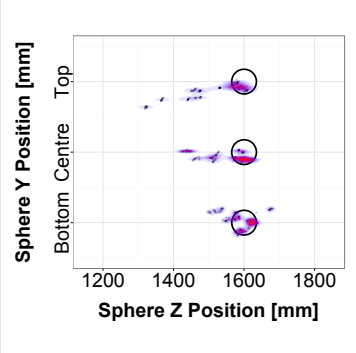
(b) 1400mm z plane - Centre positions



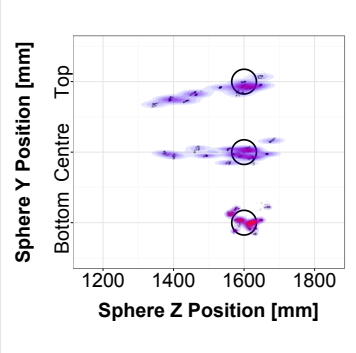
(c) 1400mm z plane - Right positions



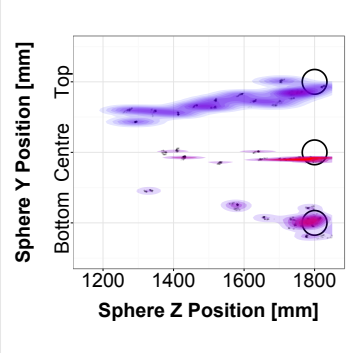
(d) 1600mm z plane - Left positions



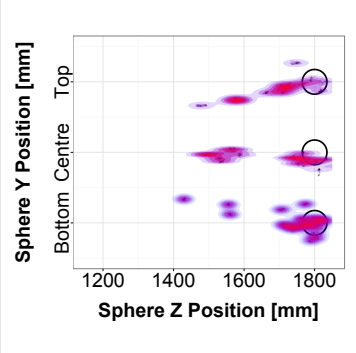
(e) 1600mm z plane - Centre positions



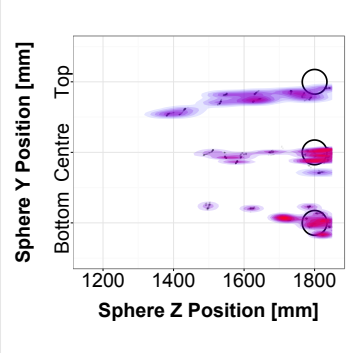
(f) 1600mm z plane - Right positions



(g) 1400mm z plane - Left positions



(h) 1400mm z plane - Centre positions



(i) 1400mm z plane - Right positions

Figure 5.11: gmp placement in the z axis of all participants for grasping spheres in 27 positions in 3 z planes (First row: 1400mm z plane. Second row: 1600mm z plane. Third row: 1800mm z plane): squares indicate sphere positions (Left, Centre and Right), and density heat maps indicate gmp placement across participants (red indicates higher density)

Table 5.9: Descriptive Statistics of $GDisp_z$ (Mean ± SD)

1400mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cube (Constant Size - 80mm)	Top	-65.70 ± 68.36	-21.72 ± 38.82	-20.03 ± 58.36
	Centre	-12.26 ± 42.97	-4.34 ± 25.51	-0.98 ± 26.33
	Bottom	0.73 ± 64.25	-1.36 ± 50.85	5.94 ± 34.00
Spheres (Constant Size - 70mm)	Top	-60.44 ± 75.96	-40.39 ± 59.44	-19.48 ± 66.01
	Centre	-12.11 ± 35.61	-6.17 ± 26.68	-1.77 ± 51.50
	Bottom	20.55 ± 43.38	11.46 ± 43.57	12.87 ± 28.38
1600mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	-129.32 ± 105.30	-98.24 ± 113.29	-53.76 ± 85.90
	Centre	-53.18 ± 86.73	-85.24 ± 104.04	-34.32 ± 83.12
	Bottom	-44.49 ± 88.57	-22.88 ± 56.20	-7.28 ± 38.87
Spheres (Constant Size - 70mm)	Top	-106.78 ± 107.01	-82.97 ± 87.37	-72.44 ± 99.93
	Centre	-85.16 ± 113.66	-49.60 ± 62.35	-37.96 ± 87.15
	Bottom	-26.07 ± 62.43	-8.20 ± 40.61	4.77 ± 30.20
1800mm Z Plane				
Object Type	Position (y)	Position (x)		
		Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	-220.12 ± 181.69	-135.87 ± 159.62	-198.07 ± 167.64
	Centre	-164.39 ± 133.31	-167.98 ± 150.09	-154.92 ± 165.29
	Bottom	-77.55 ± 119.62	-70.59 ± 130.98	-53.05 ± 120.73
Spheres (Constant Size - 70mm)	Top	-233.25 ± 174.06	-96.18 ± 106.77	-108.73 ± 158.88
	Centre	-155.76 ± 159.80	-132.85 ± 123.58	-60.61 ± 120.38
	Bottom	-81.21 ± 129.96	-78.78± 112.20	-41.74 ± 91.39

As shown in Figure 5.8j [page 83], lowest mean $GDisp_z$ was shown by users in the Bottom positions alongside the y axis across all z planes for cubes (-30.06 ± 91.07) and spheres (-20.71 ± 82.34). Similar to $GDisp_y$, bottom positions are preferred by users due to the clear visualisation of the hand and feedback monitor where the arm does no intersect the FOV of users (see Figure 5.12 [page 93]), thus leading to more accurate depth estimation. In addition, bottom positions also restricted the movements of users, thus the margin for error in bottom positions was limited as users changed their standing posture in many cases to be able to accurately grasp objects in bottom positions.

Alongside the x axis (see Figure 5.8e [page 83]), users showed the lowest $GDisp_y$ in Right positions across all z planes for cubes (-57.38 ± 121.08) and spheres (-36.12 ± 98.37) (see Table 5.9 [page 92]), similar to $GDisp_x$ and $GDisp_y$ this preference for right positions is expected as all users in this study were right handed.

Similar to Experiment 1, changing object position has also opted participants to adapt their grasp posture in terms of dexterity and type.

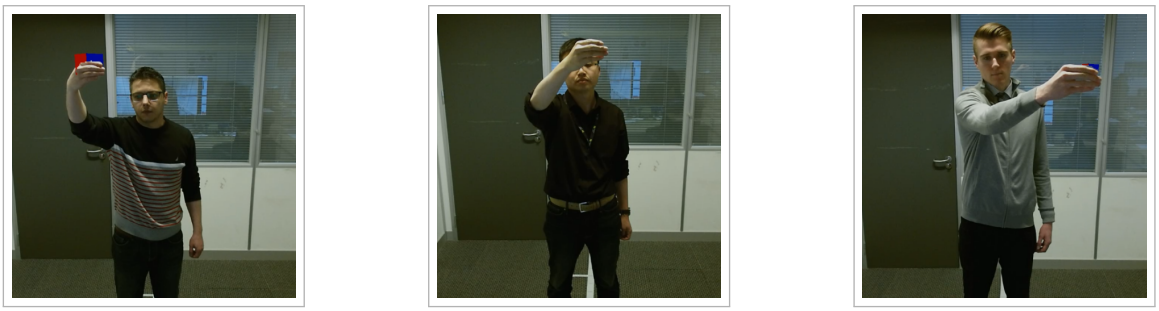


Figure 5.12: Top positions are prone to the forearm intersecting the FOV of users in the current setting, this inability to clearly visualise the full hand can hinder depth estimation

5.4.3.7 Results - Object Type

1400mm Z plane

Statistically significant differences in GAp ($\chi^2(1) = 1251$, $p < 0.01$), completion time ($\chi^2(1) = 226$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 299$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 29$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 74$, $p < 0.01$).

1600mm Z plane

Statistically significant differences in GAp ($\chi^2(1) = 636$, $p < 0.01$), completion time ($\chi^2(1) = 54$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 456$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 12$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 17$, $p < 0.01$).

1800mm Z plane

Statistically significant differences in GAp ($\chi^2(1) = 52$, $p < 0.01$), completion time ($\chi^2(1) = 357$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 31$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 41$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 102$, $p < 0.01$)

5.4.3.8 Analysis - Object Type

In this section, findings for different object types (cubes and spheres) are reported per each z plane, and not for each individual position in this experiment for clarity and to avoid repetition with results previously reported.

In GAp, users showed higher accuracy in grasping cubes than spheres in terms of matching GAp to object size across all positions, this was consistent across all z planes (see Table 5.10 [page 94]), and is opposite to the findings in experiment 1 where users showed higher accuracy in grasping spheres. Users also showed less variation in grasping cubes than spheres, thus users were more confident in size estimation of cubes using their GAp than spheres, this was consistent across all z planes with the exception of the 1800mm z plane where less variation was found for spheres than cubes (see SD values in Table 5.10 [page 94]).

In completion time, users consistently showed less mean completion times in grasping spheres than cubes across all z planes (see Table 5.10 [page 94]). Users also showed less variation in

Table 5.10: Descriptive Statistics of Object Type (Mean ± SD). Significant differences between cubes and spheres ($p < 0.01$) are marked with (*)

1400mm Z Plane					
Object Type	GAp [mm]	Completion Time [s]	$GDisp_x$ [mm]	$GDisp_y$ [mm]	$GDisp_z$ [mm]
Cubes (80mm)	81.86 ± 15.80*	6.80 ± 4.38*	27.20 ± 24.81*	-4.73 ± 17.08*	-13.30 ± 52.11*
Spheres (70mm)	74.56 ± 11.97*	5.86 ± 3.31*	23.27 ± 29.28*	-6.03 ± 21.59*	-10.61 ± 56.19*
1600mm Z Plane					
Object Type	GAp [mm]	Completion Time [s]	$GDisp_x$ [mm]	$GDisp_y$ [mm]	$GDisp_z$ [mm]
Cubes (80mm)	79.94 ± 16.17*	6.30 ± 5.29*	30.39 ± 26.90*	-9.80 ± 21.99	-58.75 ± 94.90*
Spheres (70mm)	73.87 ± 17.73*	5.14 ± 2.63*	24.32 ± 29.20*	-10.33 ± 24.13	-51.60 ± 89.06*
1800mm Z Plane					
Object Type	GAp [mm]	Completion Time [s]	$GDisp_x$ [mm]	$GDisp_y$ [mm]	$GDisp_z$ [mm]
Cubes (80mm)	76.17 ± 22.90*	6.80 ± 6.44*	30.00 ± 33.75*	-18.46 ± 36.22*	-138.06 ± 159.08*
Spheres (70mm)	73.95 ± 23.50*	5.21 ± 3.18*	31.49 ± 42.48*	-15.13 ± 39.36*	-109.90 ± 144.09

grasping spheres than cubes (see SD values in Table 5.10 [page 94]), thus users were more confident in grasping spheres as evident by completion times. Interestingly users were more accurate in grasping cubes, thus the higher mean completion times for cubes are potentially due to users spending more time on fine adjustments of their grasp.

In $GDisp_x$, users showed lower mean $GDisp_x$ in grasping spheres than cubes across all z planes, with the exception of the 1800mm z plane where a higher mean $GDisp_x$ was found for spheres than cubes (see Table 5.10 [page 94]). This is potentially attributed to the size of spheres being smaller than cubes in this experiment, thus the chance of error in gmp placement on the x axis is naturally smaller for spheres than the bigger cubes. In addition, the lack of faces in a sphere due to its shape that consists of a single surface also potentially makes it easier for users to align their gmp to the omp naturally without being constrained by faces that may influence the grasp structure and spatial position resulting in higher $GDisp_x$. This is particularly true for cubes where the visible faces, even if virtual, can hinder grasp placement in order to perform a grasp that is adjusted in accordance with the edges of the cube. For example, users grasping spheres in this study could cover the whole object with their grasp in the first attempt without having to reconstruct the posture of their grasp due to faces afterwards, whereas for cubes users naturally attempt to perform a grasp that is not only accurate, but is also naturally compliant with the geometrical features of the cube (see Figure 5.13 [page 95]). However users showed lower variation in gmp placement along the x axis consistently across all z planes in grasping cubes than spheres (see SD values in Table 5.10 [page 94]), thus users were more confident in their gmp placement along the x axis when grasping cubes than spheres, even though lower $GDisp_x$ was shown in grasping spheres.

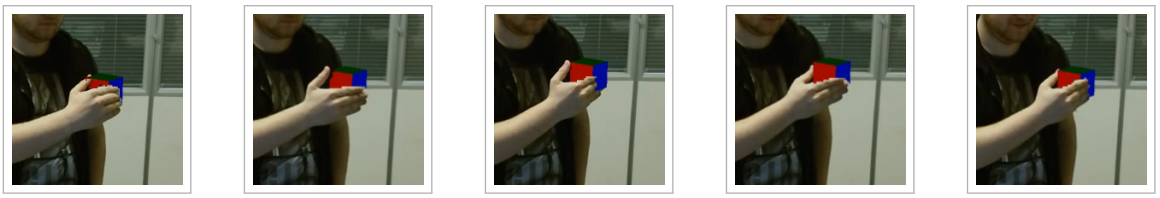


Figure 5.13: An example showing how the edges of cubes, even if virtual, can sometimes dictate the grasp structure and task completion time as users try to perform grasps that comply with the geometrical structure of the cubes presented

In $GDisp_{\text{y}}$, users consistently placed their gmp at a point lower than the omp for both objects, thus resulting in negative mean $GDisp_{\text{y}}$ across all z planes (see Table 5.10 [page 94]). Users also showed higher mean $GDisp_{\text{y}}$ in grasping spheres than cubes across all z planes with the exception of the 1800mm z plane where users showed a higher mean $GDisp_{\text{y}}$ in grasping cubes than spheres. This can again be attributed to the different sizes of both objects, where matching the gmp to the omp of spheres can be challenging to visualise due to the small size of spheres that can be fully occluded using a medium wrap grasp, whereas for cubes this problem is limited due to their bigger size that cannot be fully occluded using a medium wrap grasp, and hence users were able to better visualise and accordingly adjust their gmp in relation to the omp of the object in the y axis for cubes than spheres. This is evident by the higher variation found for spheres than cubes (see SD values in Table 5.10 [page 94]).

In $GDisp_{\text{z}}$, users consistently underestimated the position of both objects (cubes and spheres) in all z planes by placing their gmp along the z axis in a depth that is closer to the sensor than that of the virtual object (see Table 5.10 [page 94]). Users showed more accuracy in depth estimation consistently across all planes in grasping spheres than cubes, with higher variation in gmp placement along the z axis found for cubes than spheres across all planes, with the exception of the 1400mm z plane where a higher variation was found for spheres than cubes (see SD values Table 5.10 [page 94]). This can potentially again be attributed to the smaller size of spheres and their edgeless nature, however while spheres are preferable by users for depth estimation, $GDisp_{\text{z}}$ in both the experiments of this study was significantly high and users showed the lowest accuracy and highest variation in depth estimation for both objects.

Hypothesis - Revisited

$H_{1.2}$: changes in object position do not have an effect on a) grasp aperture and b) grasp displacement: **Rejected** as GAp is affected by changes in position and object type, within the bounds of the range found (66.40mm to 84.44mm), and due to large variations in gmp placement for both objects in different positions, $GDisp$ is affected by changes in size and object type.

5.5 Conclusions

This chapter presented the first baseline study to measure the accuracy of freehand grasping in exocentric AR by measuring the influence of object size, type and position on grasp performance.

Through two perceptual experiments, it was shown that GAp is constant within a range from 66mm to 88mm. This study has shown that a 70mm GAp is within the working range of participants in an AR context, and the size at which proximity of grasp aperture to object size is largest. Furthermore, it was also shown that the relationship between GAp and object size in freehand grasping is not linear as it is in grasping real objects. This is mainly due to the lack of haptic feedback, and is potentially an indication that, unlike in physical grasping in real environments, object size is not necessarily a grasp constraint that is functional in terms of determining GAp in an exocentric AR setting. This study also showed that displacement from the omp is small across the x and y axes if an object is static in position (i.e. Experiment 1), and large as position changes (i.e. Experiment 2). Notably, underestimation of object position in the z axis was found to be significantly high, even with the prior knowledge of the position of an object in the z axis (i.e. Experiment 1). This indicates that depth estimation in occlusion based freehand grasping is problematic, mainly due to the lack of user interaction awareness in the z axis and inability to fully visualise virtual objects in the z axis due to the feedback method used being a single monitor.

Based on the proposed grasp measurements presented in Chapter 4, grasping performance was found to be superior in terms of accuracy on the right and centre regions in front of participants than the left hand side. This shows that defining a working range that lies in the centre and the side of participants' dominant hand can potentially improve grasping interactions in AR. This study also demonstrated that participants adapted their grasp type and orientation according to changes in object size and position, even though one specific grasp was chosen for this work (i.e. medium wrap grasp). In addition, insights regarding user virtual object preference were noted, namely that users were faster in finishing interaction tasks when interacting with spheres but more accurate in terms of size estimation when grasping cubes. This can be attributed to the structural differences between the sphere and cube, as the lack of faces in a sphere due to its shape that consists of a single surface was potentially perceived by users as an easier object to interact with naturally without being constrained by faces or edges such as the ones present in a cube that can influence the grasp structure and spatial position. The reported presence of $GDisp_{\text{x}}$, $GDisp_{\text{y}}$ and $GDisp_{\text{z}}$ in this study showed that the proposed metric for measuring freehand grasping of virtual objects (i.e. $GDisp$) is a good indication of human behaviour in freehand grasping of virtual objects. Therefore, using $GDisp$ in parallel with GAp that is used for physical grasping to asses similarities between a grasp and the spatial and physical properties of a virtual object could provide more robust analysis of freehand grasping in AR.

In conclusion, this chapter measured the accuracy of freehand grasping in exocentric AR against

virtual object size, position and type. Findings in this study highlighted key problems in freehand grasping, namely inaccurate size estimation of virtual objects using GAp and significant overestimation of virtual object position in the z axis. These problems are largely influenced by the feedback method used in this study being visual on a single monitor in front of users. Accordingly, users had potentially limited awareness in depth judgements and object size estimation even with occlusion present. In Chapter 6 these two key problems will be addressed through adopting a novel dual view visual feedback method, to enable users to visualise their grasping interaction along the z axis thus giving users additional visual feedback to correct their grasp positioning and aperture. The next chapter will revisit the problems presented in this study by quantifying the impact of this proposed feedback methods on freehand grasping accuracy. In addition, the usability of dual view visual feedback for freehand grasping will be discussed.

Chapter 6

Study 2: Dual View Visual Feedback for Freehand Grasping

This work was published in the proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST) as "Improving freehand placement for grasping virtual objects via dual view visual feedback in mixed reality" (Al-Kalbani et al. (2016*b*))

6.1 Introduction

This chapter will present the second user study out of the four independent user studies in this work (see Figure 4.11 [page 64]) to assess the accuracy and usability of freehand grasping in exocentric AR. This study is a replication of Study 1 (Chapter 5) with the addition of a second visual view of interaction. Findings in Chapter 5 highlighted the key problems associated with freehand grasping in exocentric AR using the grasp metrics proposed in this thesis (GAp and $GDisp$), namely how users often fail to accurately estimate the correct depth location of virtual objects and how the grasp aperture does not change linearly to the changes in virtual object size. These two problems in Chapter 5 were highly influenced by the feedback method used being single view visual feedback. This form of feedback did not allow users to be fully aware of the spatial position of their grasp in relation to a virtual object, especially in the z axis where grasp displacement was highest. In addition users also showed high variation in grasp aperture and structure due to their inability to visualise their full hand during interaction using single view visual feedback. This chapter will aim to address these problems through the feedback method used.

Feedback is defined in a general context as the process in which the impact of an action is returned to improve or correct the next action. Absence of feedback can lead to poorer performance in AR environments (Maria et al., 2015) and use of suitable feedback can lead to direct improvements in user performance (Pitts et al., 2012). Feedback modalities vary, with visual, audio, haptic, tactile and force feedback commonly used within AR. Multimodal feedback is also

widely used in current research, and aims to improve user performance through the integration of two or more feedback modalities, thus giving users a higher sense of presence in relation to virtual elements in AR environments. For example Duff et al. (2010) presented a mixed reality system for stroke rehabilitation with multimodal visual and aural feedback. Vieira et al. (2015) also used haptic and audio feedback modalities alongside projection mapping techniques for visual feedback, to increase sources of awareness in rehabilitation tasks. Current research shows that methods combining different feedback modalities with visual feedback do benefit performance in interaction, however this can be limited in this work where freehand grasping, without any wearable device, is required and visual feedback that is considered to be a conventional type of feedback (Prattichizzo et al., 2012) in AR is commonly used alone.

This chapter will enhance the visual feedback method used in this work to address the problems discussed in Chapter 5, through implementing a novel dual view visual feedback method for assisting freehand grasping of virtual objects in two separate experiments. Results in this chapter will be directly compared to the results found in Chapter 5 to measure the impact of this proposed dual view visual feedback method in comparison to single view visual feedback that was used in Chapter 5. The impact of this proposed visual feedback method on freehand grasping accuracy will be measured using GAp and $GDisp$, and the usability of this proposed feedback method will also be addressed using the standardised System Usability Scale (SUS). Section 6.2 [page 99] firstly outlines the design of the two experiments in this study in terms of the conditions under test, participants recruited, statistical model used and experiment protocol. Section 6.3 [page 104] then discusses the data collected in the first experiment of this study, and provides a comprehensive analysis of the interaction and usability results. This is followed by a discussion of the second experiment in this study and a comprehensive analysis of the interaction and usability results in Section 6.4 [page 113]. Finally Section 6.5 [page 135] provides the conclusions drawn from this study and a summary of the key outputs that will be addressed in following chapters.

6.2 Study Outline

6.2.1 Design

Two experiments were conducted in this study using the baseline setup detailed in Chapter 4 with the addition of a Live! Cam Optia Pro HD webcam[1] to provide additional visual feedback (see Figure 6.1 [page 101]):

- Experiment 1 to quantify the influence of object size and object type on grasp accuracy using dual view visual feedback
- Experiment 2 to test the influence of object position and object type in x,y and z space on grasp accuracy using dual view visual feedback

[1] http://support.creative.com/kb/ShowArticle.aspx?sid=10859

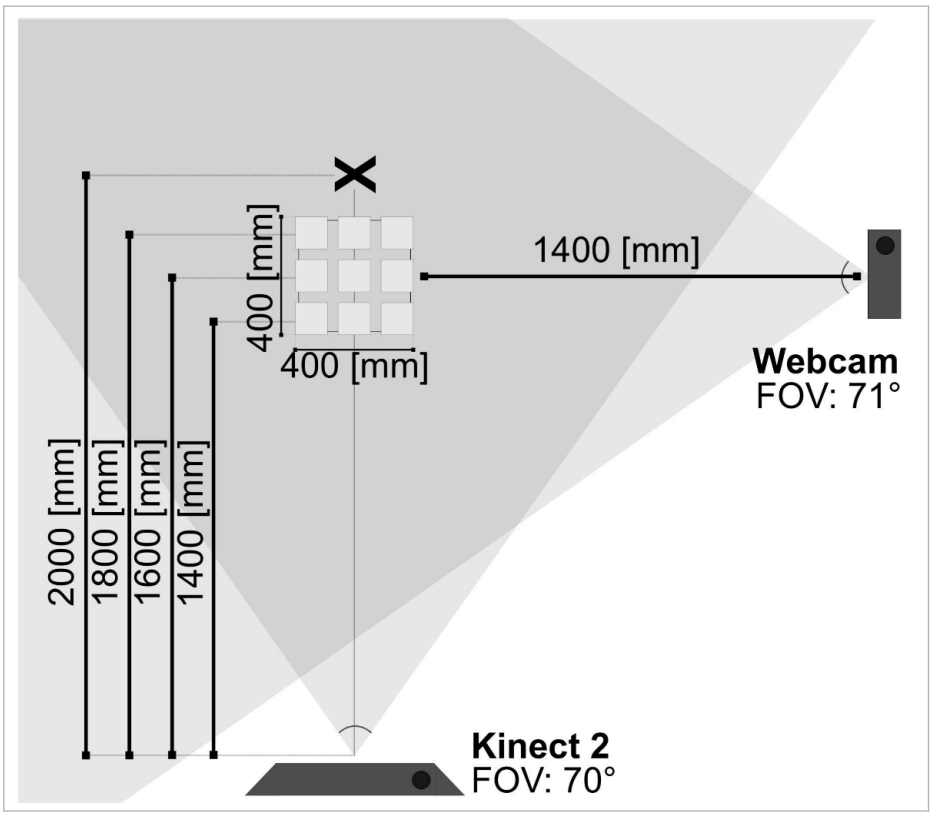
An HD webcam (FOV: 71.0°) is used as a second view feedback camera. FOV of the Kinect and webcam were comparable and the full interaction space was visible on both views. Sensor placement and standing position of participants away from the sensor were identical to the one outlined in the baseline setup (see Section 4.2.3.1 [page 54].

Second View Visual Feedback Configuration

Given that this study is a replication of study of Study 1 (see 5 66, it was essential for this study to follow the same physical configuration and environment conditions in order to be able to directly compare the results between the two feedback methods (i.e. single view in Chapter 6 and dual view visual feedback in this study). For this reason, placement of the additional side view camera had to be integrated within the baseline environment used in Chapter 5 without interfering with user or sensor performance. Placing the additional camera either at the top or to the side of users provide this integration of the additional camera in the current environment. After considering these two options, it became apparent that placing the camera at the top of users to provide an aerial view of their interaction in the z axis could result in a flat rendering of virtual objects and can potentially impact performance (i.e. cubes would look like squares and spheres like circles), for this reason positioning the additional camera to the side of users was implemented instead.

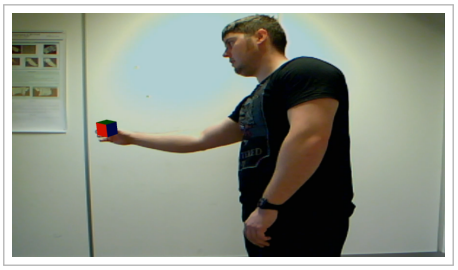
For the additional view in this study, participants stood 1400mm away from the side view webcam (see Figure 6.1a [page 101]), placed to the left hand side and at the same height as the Centre Middle position (1250mm) presented to participants in Object Position Experiment (see Table 6.1 [page 102]). This was done to ensure all objects in varying positions in the Object Position are visible to participants on the feedback monitor. As the distance to the webcam was smaller to the one from the Kinect sensor, 3D virtual objects were computed to be larger in OpenGL to reflect an accurate representation of the closer distance to participants and this scales comparably to the user's hand and body.

Second view visual feedback was placed to the side of participants as results in Chapter 5 have shown that Grasp Displacement in the x axis was user dependent and was influenced by the dominant hand of users, not the feedback method (see Figure 6.1b [page 101]). On the other hand, Grasp Displacement in the y axis was influenced by the feedback method, thus spatial placement of the hand in the y axis was affected by the visual feedback method used. Moreover, highest Grasp displacement was found in the z axis due to using single view visual feedback, thus a side view as a second visual feedback method is used to show the y and z axes (see Figure 6.1c [page 101]), the two axes that were directly affected by the feedback method used in Chapter 5. Moreover, high grasp variation was also found in Chapter 5, meaning that participants used different grasp types to the one they were instructed to use in the test. This behaviour was attributed to participants trying to visualise their full hand. Adding a side view allows participants to visualise all parts of their hand without the need to adapt their grasp type. The feedback monitor was split into two equally sized side by side windows, showing the frontal view feedback from the sensor on the left hand side window, and the side view feedback from



(a) Experiment Setup

(b) Front View



(c) Side View

(d) Dual View Visual Feedback

Figure 6.1: Setup of the dual view visual feedback system developed

the webcam on right hand side window (see Figure 6.1d [page 101]). Positions of the windows on the feedback monitor were unchanged throughout the study. However, participants were asked to comment on the positions of the windows and their influence on their performance in the subjective analysis after the experiments.

Conditions of both experiments are shown in Table 6.1[page 102], with the accuracy of a medium wrap grasp measured against the proposed metrics in this thesis; grasp aperture (GAp) and grasp displacement ($GDisp$) using dual view visual feedback. To represent the accuracy of a grasp independent of additional rendering, for both experiments, the baseline objects which have not undergone complex rendering and represent a simple abstract shape are used.

Table 6.1: Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor

<table>
<tr><th colspan="5">Experiment 1 - Object Size</th></tr>
<tr><td>Condition</td><td colspan="4">Levels</td></tr>
<tr><td>Object Size [mm]</td><td colspan="4">40 - 50 - 60 - 70 - 80 - 100</td></tr>
<tr><td>Object Type</td><td colspan="4">Cube and Sphere</td></tr>
<tr><th colspan="5">Experiment 2 - Object Position</th></tr>
<tr><td>Condition</td><td colspan="4">Levels</td></tr>
<tr><td rowspan="5">Object Position (x, y) [mm]</td><td></td><td>Left</td><td>Centre</td><td>Right</td></tr>
<tr><td>Top</td><td>-400, 1650</td><td>0, 1650</td><td>400, 1650</td></tr>
<tr><td>Centre</td><td>-400, 1250</td><td>0, 1250</td><td>400, 1250</td></tr>
<tr><td>Bottom</td><td>-400, 850</td><td>0, 850</td><td>400, 850</td></tr>
<tr><td colspan="4">* 9 positions were repeated in each z plane (1400mm - 1600mm - 1800mm), resulting in 27 positions in total</td></tr>
<tr><td>Object Type</td><td colspan="4">Cube and Sphere</td></tr>
</table>

Hypotheses

$H_{2.1}$: using dual visual feedback in grasping virtual objects that change in size has no effect on a) grasp aperture and b) grasp displacement (Experiment 1).

$H_{2.2}$: using dual visual feedback in grasping virtual objects that change in position has no effect on a) grasp aperture and b) grasp displacement (Experiment 2).

6.2.2 Participants

30 participants ranged in age from 21 to 62 (M = 31.77, SD = 10.64), in arm length from 480mm to 660mm (M = 566.00, SD = 41.49), in hand size from 170mm to 200mm (M = 187.45, SD = 9.97), in height from 1570mm to 1950mm (M = 1761.36, SD = 95.03) and 7 were female and 23 male.

Taking into account balance in hand size, arm length, gender, age and height, participants were separated into two groups of 15 for each experiment.

6.2.3 Statistical Analysis

Statistical models used in the analysis of results in this study were validated using assumptions of different models. Kruskal Wallis H test (Kruskal and Wallis, 1952), a rank-based non-parametric test, is used to analyse the data collected in this study. Statistical significance of the Kruskal Wallis H test results is implemented using a post-hoc test for multiple comparisons using Dunn Test with Bonferroni correction (Dunn, 1961).

6.2.4 Protocol

This study followed the baseline experiment protocol outlined in Section 4.2.4.2[page 62] prior to collection of data.
Participants underwent initial training of the medium wrap grasp on real and virtual objects and were given time to familiarise themselves with the side view visual feedback concept. The test coordinator explained the procedure between each block of tests (i.e cube and sphere), and participants were allowed to rest before the presentation of every object. Each experiment was formed of a 5 minutes training/instruction session, 10 minutes of grasping a cuboid object, 5 minutes break and 10 minutes of grasping a spherical object (order of virtual objects counter-balanced).
After completing the test, participants were asked to fill in a usability questionnaire and a set of questions regarding their interaction with the system. The usability of the system was evaluated by a user satisfaction test based on the System Usability Scale (SUS) (Brooke, 1996). This questionnaire consists of 10 items, which were evaluated by using a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Through feedback from this questionnaire, the ease of use and usability of this new configuration of the system is evaluated.
In order to further assess interaction strategies and behaviour protocols by participants while using the system, they were asked to answer a set of 5 close-ended questions. These questions were presented as a post-test questionnaire and participants commented on anything they considered related to their interaction and the system (see Appendix A [page 211]). Questions were:

1. Which screen did you look at first?
2. Which screen did you depend on the most?
3. Which view did you find to be more important?
4. Did you use the dual view in a specific order?
5. Do you think changing positions of both feedback screens would make a difference in performance?

6.3 Experiment 1 - Object Size

6.3.1 Experiment Design

A 2 × 6 within-subjects design was used, with two primary conditions: object size and object type (see Table 6.1 [page 102]). All 15 participants took part in both conditions. Every permutation for both object types was randomly presented to participants to exclude potential learning effects. In total, each participant completed 6 (sizes) × 5 (repetitions) × 2 (objects) = 60 trials and 900 grasps in total (60 trials × 15 participants). Each static grasp of every participant was recorded for 5 seconds (75 frames), leading to collecting 67500 raw data points (900 grasps × 75 frames).

6.3.2 Procedure

For this first experiment, participants were instructed to accurately match their grasp aperture to the size and position of the virtual object in the shortest time possible on both feedback views. Before interaction, an object (cube or sphere) appeared on the feedback monitor, in 6 different sizes (see Table 6.1 [page 102]). Objects were positioned 1600mm away from the sensor and 400mm away from participants (z), at a height of 1250mm (y) and at the zero (x) point on the sensor. This position was constant throughout the experiment.
During the experiment, all participants were instructed to verbally inform the test coordinator that they are satisfied with the grasp they have performed on both feedback views (front and side), and maintain the grasp for 5 seconds while the measurements are stored.

6.3.3 Results

6.3.3.1 Results - Grasp Aperture (GAp)

A statistically significant difference was found in Grasp Aperture (GAp) between the two visual feedback methods (single view and dual view) in grasping spheres ($\chi^2(1) = 1270.90$, $p < 0.01$) and cubes ($\chi^2(1) = 5.06$, $p < 0.01$).

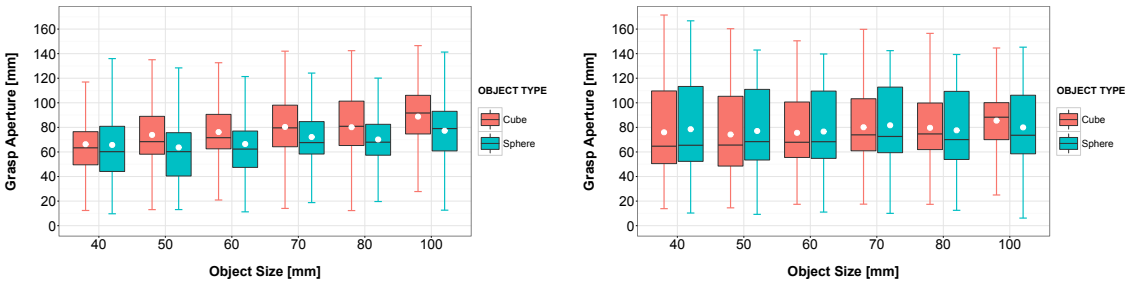
6.3.3.2 Analysis - Grasp Aperture (GAp)

As shown in Table 6.2 [page 105], participants maintained their behaviour in matching their GAp to object size with the addition of side view visual feedback where overestimation of object size occurred in 17 out of the total 24 trials in this experiment in both the single and dual view conditions for both objects. For both objects, participants overestimated object size up until the size that had the lowest mean difference between GAp and object size (80mm for cubes and spheres). In addition, both objects showed that with the 100mm size, participants underestimated its size by a mean of -14.51mm for cubes (SD = 24.92), and -20.02mm for spheres (SD = 28.59). This behaviour was present in both conditions (single and dual view visual feedback).

Table 6.2: Descriptive Statistics of GAp, $GDisp_x$, $GDisp_y$, $GDisp_z$ and completion time for different sizes of cubes and spheres (Mean ± SD). Statistical significance ($p < 0.01$) between between single and dual view visual feedback methods are marked with (*)

Object Type	Object Size	View	GAp	$GDisp_x$	$GDisp_y$	$GDisp_z$	Time
Cubes	40	Single	66.31 ± 29.97*	31.45 ± 14.10*	-15.81 ± 12.15*	-34.34 ± 65.58*	4.28 ± 2.05*
		Dual	76.01 ± 33.94*	23.61 ± 23.07*	-15.33 ± 11.44*	-6.29 ± 24.86*	7.84 ± 3.64
	50	Single	73.89 ± 28.16*	26.83 ± 13.33*	-14.76 ± 11.93*	-38.75 ± 60.73*	4.43 ± 1.99*
		Dual	74.21 ± 32.53*	20.53 ± 20.85*	-14.04 ± 10.82*	-5.09 ± 29.53*	6.99 ± 3.84*
	60	Single	76.18 ± 24.12*	26.09 ± 14.18*	-13.17 ± 12.57	-40.37 ± 62.10*	4.16 ± 1.90*
		Dual	75.57 ± 31.28*	17.42 ± 19.67*	-12.93 ± 10.01	-8.56 ± 20.22*	7.04 ± 3.10*
	70	Single	80.38 ± 22.55*	25.35 ± 14.75*	-11.15 ± 10.37	-42.13 ± 55.54*	4.25 ± 1.61*
		Dual	80.10 ± 30.03*	17.88 ± 21.57*	-11.65 ± 10.95	-6.31 ± 19.66*	7.27 ± 4.05*
	80	Single	80.13 ± 24.59*	29.40 ± 14.69*	-10.53 ± 10.62*	-35.31 ± 56.68*	3.96 ± 2.03*
		Dual	79.67 ± 28.57*	19.38 ± 21.41*	-15.13 ± 11.11*	-12.56 ± 20.07*	7.21 ± 3.56*
	100	Single	88.77 ± 22.39*	28.93 ± 12.31*	-8.90 ± 12.36*	-39.42 ± 68.15*	4.48 ± 1.96*
		Dual	85.49 ± 24.92*	17.15 ± 20.99*	-8.09 ± 10.53*	-14.96 ± 17.00*	9.57 ± 10.83*
Spheres	40	Single	65.73 ± 30.83 *	36.51 ± 13.31*	-12.62 ± 12.76*	-29.21 ± 60.92*	3.89 ± 1.67*
		Dual	78.51 ± 33.70*	38.14 ± 19.57*	-6.42 ± 12.29*	6.93 ± 27.92*	6.49 ± 3.46*
	50	Single	63.71 ± 30.4*	36.10 ± 13.91*	-12.82 ± 13.03*	-31.86 ± 62.98*	3.57 ± 1.52*
		Dual	77.08 ± 33.74*	37.78 ± 20.24*	-9.43 ± 11.70*	0.05 ± 31.59*	6.05 ± 3.42*
	60	Single	66.48 ± 28.97*	31.98 ± 14.96*	-11.97 ± 11.90*	-29.44 ± 57.61*	3.45 ± 1.37*
		Dual	76.6 ± 32.54*	28.56 ± 19.78*	-4.83 ± 12.31*	0.61 ± 31.22*	5.80 ± 2.63*
	70	Single	72.18 ± 26.03*	30.41 ± 14.52*	-8.41 ± 12.08*	-27.04 ± 61.96*	3.33 ± 1.33*
		Dual	81.71 ± 31.09*	33.34 ± 20.69*	-3.45 ± 12.71*	2.53 ± 27.45*	5.88 ± 3.10*
	80	Single	70.16 ± 26.20*	29.53 ± 13.59*	-7.46 ± 12.89*	-31.08 ± 62.04*	3.23 ± 1.29*
		Dual	77.61 ± 31.53*	24.69 ± 17.66*	-3.58 ± 10.86*	-0.24 ± 24.04*	5.84 ± 3.14*
	100	Single	77.24 ± 24.35	24.56 ± 14.26*	-5.77 ± 10.39*	-30.62 ± 57.20*	3.41 ± 1.95*
		Dual	79.98 ± 28.59	23.58 ± 18.49*	-1.41 ± 10.47*	-2.69 ± 20.46*	6.43 ± 3.21

Users showed higher accuracy in size estimation using their GAp in the single view condition, where the additional dual view outperformed single view visual feedback in size estimation in just 4 out of the 24 trials in this experiment (see Table 6.2 [page 105]). In addition, as shown in Figure 6.2b [page 106] users also showed higher variation in their GAp under the dual view visual feedback condition in every trial in this experiment (see Table 6.2 [page 105]). Users also showed a narrower mean GAp range across all sizes and object types in the dual view visual feedback condition (from 74.21 ± 32.53 to 85.49 ± 24.92) than the range of mean GAp reported for single view visual feedback (from 65.73 ± 30.83 to 88.77 ± 22.39) (see Figure 6.2a). Given that object sizes ranged from 40mm to 100mm, this shows that responsiveness of participants in terms of accurately matching GAp to object size is constrained between 60mm and 80mm, regardless of the feedback method used, thus again showing that, unlike grasping real objects, freehand grasping is not dictated by object size as a grasp constraint in exocentric AR. This finding was surprising in this experiment as participants had an additional side view visual feedback that clearly showed their thumb and index finger to accurately match their GAp to object



(a) GAp across different object sizes - Single View Visual Feedback [from Study 1 - Chapter 5]

(b) GAp across different object sizes - Dual View Visual Feedback

Figure 6.2: GAp for different object sizes in the 1600mm z plane using: 6.2a [page 106] Single view visual feedback and 6.2b [page 106] Dual view visual visual feedback. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

size. However dual visual feedback did not show any improvements over single view feedback in GAp matching to object size potentially due to the additional cognitive load associated with dual view visual feedback, where users were instructed to adjust their grasp using two separate views simultaneously.

6.3.3.3 Results - Completion Time

Statistically significant difference in completion time between the two feedback methods was found for cubes ($\chi^2(1) = 18863$, $p < 0.01$) and spheres ($\chi^2(1) = 16551$, $p < 0.01$).

6.3.3.4 Analysis - Completion Time

Dual view visual feedback significantly increased completion times in comparison to single view visual feedback in all trials in this experiment (see Table 6.2 [page 105]), where an increase in overall completion times across all sizes reported for single view visual feedback (4.26s $\pm$ 1.93 for cubes and 3.48s $\pm$ 1.55 for spheres) was found for dual view visual feedback for both cubes (7.65s $\pm$ 5.61) and spheres (6.08s $\pm$ 3.18). This was expected as adding a side view camera for dual visual feedback makes participants aware of their inaccuracy in grasp placement, and leads participants to spend more time fine adjusting their grasp for the purpose of achieving more grasp accuracy. However, this does not necessarily lead to better accuracy in size estimation using GAp.

6.3.3.5 Results - Grasp Displacement ($GDisp$)

Statistically significant differences were found between the two visual feedback methods in $GDisp_x$(for cubes ($\chi^2(1) = 2875.70$, $p < 0.01$), but not spheres ($\chi^2(1) = 4.20$, $p > 0.01$), $GDisp_y$ (for spheres ($\chi^2(1) = 2551.50$, $p < 0.01$), but not cubes ($\chi^2(1) = 5.89$, $p > 0.01$) and $GDisp_z$ (for cubes ($\chi^2(1) = 2420.30$, $p < 0.01$) and spheres ($\chi^2(1) = 5752.40$, $p < 0.01$).

6.3.3.6 Analysis - Grasp Displacement ($GDisp$)

***GDisp*$_{\text{x}}$**

Similar to single view visual feedback, positive $GDisp_x$ was present for both objects. This positive $GDisp_x$ is expected as all participants were right-handed, and the Grasp Middle Point (gmp) was computed on the right hand side of virtual objects. Dual view visual feedback reduced mean $GDisp_{\text{x}}$ in all trials in this experiment (see Table 6.2 [page 105]), and mean $GDisp_x$ was lower for both objects across all sizes using dual view visual feedback (19.33mm $\pm$ 21.40 for cubes, and 31.01mm $\pm$ 20.28 for spheres) than single view visual feedback (28.01mm $\pm$ 14.08 for cubes and 31.52 $\pm$ 14.68 for spheres). This shows that adding side view visual feedback significantly improves the gmp spatial positioning in the x axis and reduces $GDisp_x$ from the centre of virtual objects.

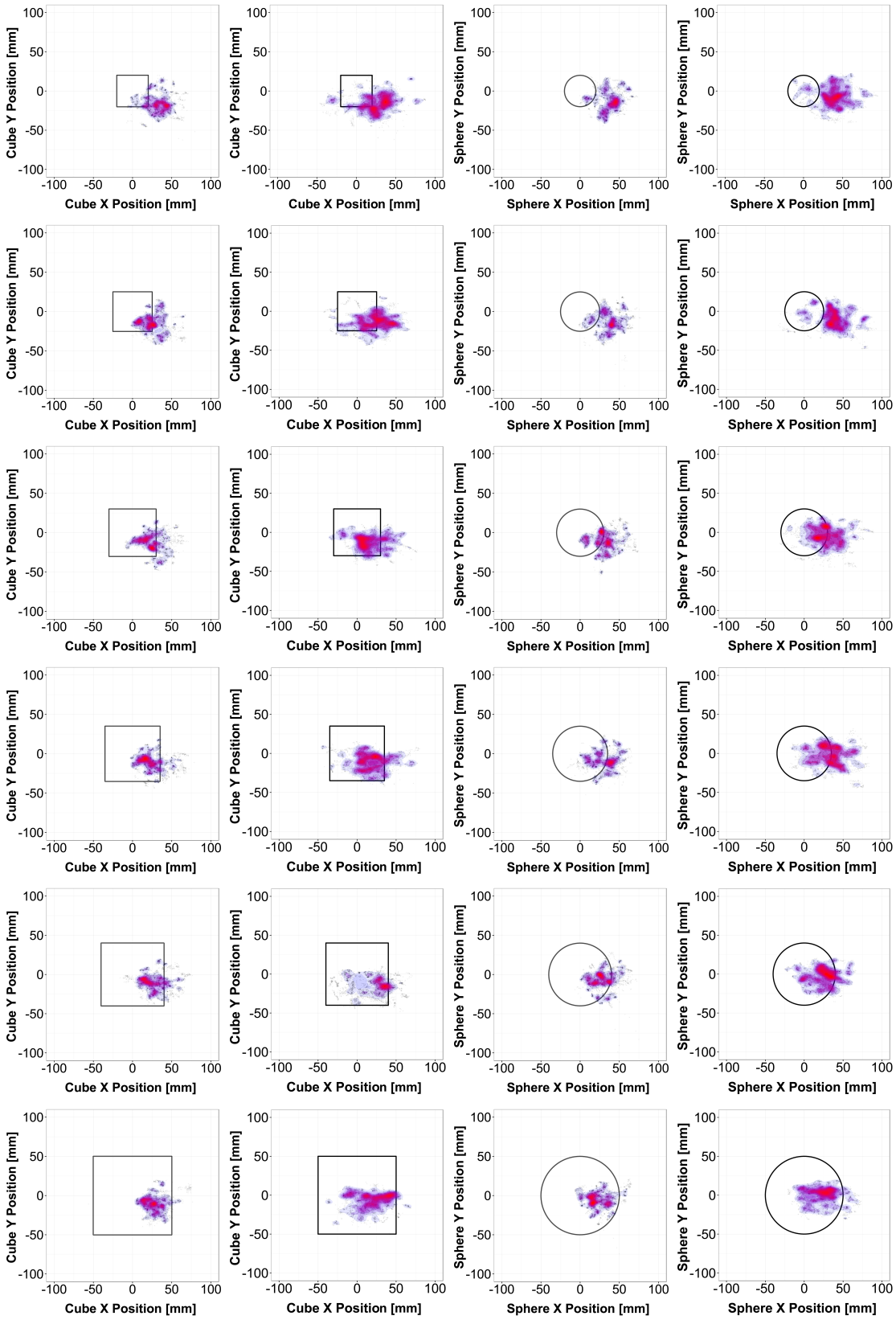
Similarity in gmp placement on the x axis was found in the first study in Chapter 5 for single view visual feedback as shown by the range of clusters on the x axis in Figures 6.3a [page 108] and 6.3c [page 108]. Mean $GDisp_x$ across all sizes of cubes and spheres is reduced using dual view visual feedback thus accordingly shifting gmp placement of users along the x axis more towards the omp as shown in Figures 6.3b [page 108] and 6.3d [page 108].

***GDisp*$_{\text{y}}$**

Negative $GDisp_y$ was present for both objects in all trials in this experiment (see Table 6.2 [page 105]). This reveals that participants placed their gmp below the Object Middle Point (omp), a behaviour that was also present in single view feedback and is potentially attributed to participants trying to show parts of the objects presented to them on the feedback monitor, a strategy that reassured participants that they have grasped the virtual object.

Users showed lower mean $GDisp_y$ in 10 out of the 12 sizes in this experiment using dual view visual feedback (see Table 6.2 [page 105]), where mean $GDisp_y$ was lower for both objects across all sizes using dual view visual feedback (-12.13mm $\pm$ 11.10 for cubes and -4.85mm $\pm$ = 12.02 for spheres) than single view visual feedback (-12.37mm $\pm$ 11.94 for cubes and -9.84mm $\pm$ 12.51 for spheres). This shows that dual view visual feedback significantly improves the gmp spatial positioning in the y axis by reducing $GDisp_y$ from the centre of virtual objects. Similar to $GDisp_x$, gmp placement across participants on the y axis was comparable across object sizes as shown by the range of clusters in Figures 6.3b [page 108] and 6.3d [page 108]. This consistency in gmp placement on the y axis was also present using single view visual feedback (see Figures 6.3a [page 108] and 6.3c [page 108]).

Mean $GDisp_x$ and $GDisp_y$ for each object size in both objects have shown that placement of gmp shifted towards the 0 origin of the x and y axes as shown in Figures 6.3b [page 108] and 6.3d [page 108], this indicates that even though $GDisp_x$ and $GDisp_y$ are still existent with the use of dual view visual feedback, the displacement is reduced and is closer to the origin of the virtual object than it was with using single view visual feedback. Moreover, SD differences of $GDisp_x$ and $GDisp_y$ means within object sizes between cubes and spheres were comparable, indicating



(a) Cubes - Single View (b) Cubes - Dual View (c) Spheres - Single View (d) Spheres - Dual View

Figure 6.3: gmp placement in the x and y axes for cubes and spheres of all participants in six sizes (40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 6.3a and 6.3c: Single view visual feedback. 6.3b and 6.3d: Dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density)

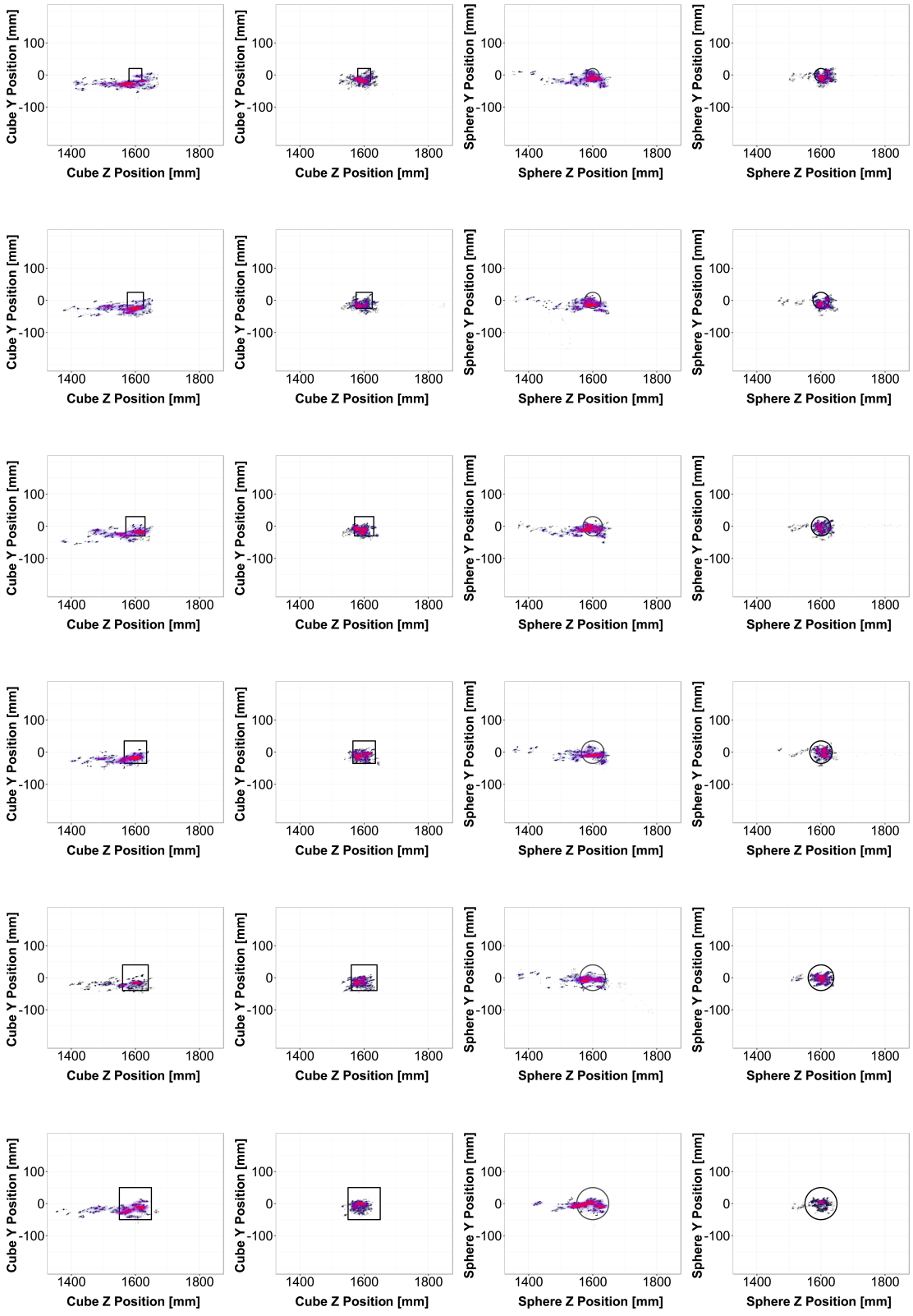
that contact of gmp with the surface of the object was reflective of size growth of objects rather than movements by participants. This behaviour was again present in the two conditions (single and dual view visual feedback), and it shows that even though dual view visual feedback reduces $GDisp_x$ and $GDisp_y$ and moves participants closer to the centroid of virtual objects in the x and y axis, participants remain consistent in their spatial gmp placement regardless of changes in object size. This consistency is expected as object position was unchanged throughout this experiment.

$GDisp_z$

$GDisp_z$ presented the highest displacement out of all three axes with single view visual feedback in Study 1 (Chapter 5). In this study, whether dual view visual feedback can mitigate high $GDisp_z$ and aid in achieving accurate depth positioning in AR is tested.

Negative mean $GDisp_z$ was found for both objects across all sizes (see Table 6.2 [page 105]), this indicates that majority of participants underestimated the z position of omp by placing their gmp in front of the omp for all sizes. Overestimation of z position was also present, but not as frequent as underestimation, as 54% of the data showed underestimation, while overestimation was present in 45% of the data. However, the difference between overestimation and underestimation is smaller and distributed in a more balanced manner when using dual view visual feedback than single view visual feedback, as underestimation was found to be present in 67% of the data, and overestimation was present in 33% of the data using single view visual feedback. Position of gmp in the z axis was comparable across all sizes for both objects, and more clustered in the centre of objects as shown by Figures 6.4b [page 110] and 6.4d [page 110]. This is attributed to the more balanced distribution of z position overestimation and underestimation caused by dual view visual feedback, and the constant position of virtual objects in this experiment.

Users showed lower mean $GDisp_z$ in all trials in this experiment under the dual view visual feedback condition (see Table 6.2 [page 105]), where mean $GDisp_z$ was lower using dual view visual feedback (-8.96mm $\pm$ 22.56 for cubes and 1.20mm $\pm$ 27.55 for spheres) as shown in Figures 6.4b [page 110] and 6.4d [page 110] than single view visual feedback (-38.39mm $\pm$ 61.67 for cubes and -29.87mm $\pm$ 60.51 for spheres). This shows that dual view visual feedback significantly reduces $GDisp_z$, and improves gmp spatial positioning in the z axis by reducing $GDisp_z$ from the centre of virtual objects thus bringing closer the gmp of users to the omp along the z axis. Moreover, dual view visual feedback reduced variation in $GDisp_z$ as shown by the SD values in Table 6.2 [page 105] when compared to the values found for single view visual feedback in Chapter 5 (see Figures 6.4a [page 110] and 6.4c [page 110]), thus showing that users were more confident in their gmp placement along the z axis when they were provided with a second view that allows visualisation of the hand in the z axis.



(a) Cubes - Single View (b) Cubes - Dual View (c) Spheres - Single View (d) Spheres - Dual View

Figure 6.4: gmp placement in the z axis for cubes and spheres of all participants in six sizes (40mm - 50mm - 60mm - 70mm - 80mm - 100mm). 6.4a and 6.4c: Single view visual feedback. 6.4b and 6.4d: Dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density)

6.3.3.7 Results - Object Type

In the dual view visual feedback condition, statistically significant differences between different object types in different sizes were found in completion time ($\chi^2(1) = 3046$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 4566$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 6967$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 5589$, $p < 0.01$). No significant differences between different object types in different sizes were found in GAp ($\chi^2(1) = 0.02$, $p > 0.01$).

6.3.3.8 Analysis - Object Type

In this section, findings for different object types (cubes and spheres) are reported for all object sizes, and not for each individual size in this experiment to avoid repetition with results previously reported.

In GAp, users showed a lower mean overestimation of object size (i.e. more accurate) in grasping cubes than spheres in every size under test in this experiment. Lower GAp variation was also found for grasping cubes than spheres in every size with the exception of the 40mm size where users showed lower GAp variation for grasping spheres than cubes (see Table 6.2 [page 105]). Participants also showed a lower mean GAp in grasping cubes (78.51mm $\pm$ 30.58) than spheres (78.58mm $\pm$ 31.96). These findings contradict those found for single view visual feedback in Chapter 5 where it was found that users were more accurate in grasping spheres than cubes. However GAp differences between the two object types in this study were not significant, thus users showed comparable GAp for both objects and generally presented a GAp working range from 74mm to 85mm regardless of object size or type.

In completion time, participants showed lower mean completion times in grasping spheres (6.08s $\pm$ 3.18) than cubes (7.65s $\pm$ 5.61) across all sizes, where users showed a lower mean completion time in in all six sizes under test in this experiment when grasping spheres (see Table 6.2 [page 105]). Dual view visual feedback expectedly increased mean completion time significantly for both objects, and users showed the same trend under the single view feedback condition where less mean completion time was also reported for spheres than cubes. However under the dual view visual feedback condition in this study, users were more accurate in size matching of cubes than spheres, thus unlike freehand grasping using single view visual feedback, longer completion times lead to higher accuracy in size matching using GAp in dual view visual feedback in this study.

In $GDisp_x$, similar to single view visual feedback, participants showed a lower mean $GDisp_x$ in grasping cubes (19.33mm $\pm$ 21.40) than sphere (31.01mm $\pm$ 20.28) across all sizes using dual view visual feedback, where users showed a lower mean $GDisp_x$ in grasping cubes across all object sizes under test in this experiment (see Table 6.2 [page 105]). Dual view visual feedback significantly reduced mean $GDisp_x$ for both objects, and the higher accuracy in terms of gmp placement along the x axis for cubes is potentially attributed to the edges of the cube that can

aid users to better place their grasp on the x axis using dual view feedback more so than spheres that lack edges.

In $GDisp_{\mathrm{y}}$, participants showed a lower mean $GDisp_{\mathrm{y}}$ in grasping spheres (-4.85mm $\pm$ 12.02) than cubes (-12.86mm $\pm$ 11.10) across all sizes, where users showed lower mean $GDisp_{\mathrm{y}}$ in all six sizes under test when grasping spheres (see Table 6.2 [page 105]), this was also found for single view visual feedback. Dual view visual feedback improved gmp placement along the y axis for both objects and significantly reduced mean $GDisp_{\mathrm{y}}$, thus users presented comparable $GDisp_{\mathrm{y}}$ for both objects as shown by the clusters in Figures 6.3b [page 108] and 6.3d [page 108].

In $GDisp_{\mathrm{z}}$, participants showed a lower mean $GDisp_{\mathrm{z}}$ in grasping spheres (1.20mm $\pm$ 27.55) than cubes (-8.96mm $\pm$ 22.56) across all sizes, where users showed lower mean $GDisp_{\mathrm{z}}$ in all six sizes under test when grasping spheres. Despite the superior depth estimation in grasping spheres, users showed significantly improved depth perception for both objects using dual view visual feedback in comparison to single view visual feedback (see Table 6.2 [page 105]).

6.3.3.9 Usability Analysis

SUS average score for the first experiment in this study was 77 (SD = 16.45). According to the SUS ranking system of Bangor et al. (2009*a*) this rating of the dual view visual feedback is "good and acceptable" in experiment 1. User comments using post test questionnaires below provide general subjective insights regarding their experience in grasping virtual objects using dual view visual feedback, however these insights may not be directly representative of user performance and accuracy during interaction as these subjective responses were not measured against performance in this work.

Out of 15 participants, 6 (37.50%) preferred to look first to the frontal view while 8 (53.33%) focused their attention on the side view first, one user remained undecided. A user that looked at the side view first commented saying "mostly as the other view (i.e. front view) does not provide depth information", and another also commented saying "the thing that varied was the size of the object which meant I had to adjust my grasp size, which required the side view". This potentially shows that users used the side view for the two issues it was aimed at solving, namely inaccurate depth and size estimation.

To the question of which view was the most important for them, the opinion was divided into 7 (46.66%) users referring to use the frontal view more, while the remaining 8 relied more on the side view (53.33%). With respect to which view was considered more important during the performance of the experiment, 7 (46.66%) users considered it to be the frontal view while 7 (46.66%) chose the side view. One user remained undecided. A user that relied more on the side view commented saying "with only the front view, you would not know if you were getting the right depth of the object accurately". Another user suggested that both views were equally important and commented that "the frontal view screen can help me to locate the object first, and the second side view can help me to keep my hand on the object stably". This again high-

lights the importance of the side view for depth judgements, and also highlights the individual differences between users in their view preference during freehand grasping.
On using the system again, 9 users (60.0%) will interact with the system again with dual visual feedback. One user suggested that using dual view visual feedback can provide more information regarding the grasping interaction and commented "I have to adjust more because of the side view, which will give more details for interaction", whereas another user pointed out that a secondary side view is only required if "depth is important" in an interaction task. 13 participants out of the 15 available had a specific approach for using dual visual feedback. The majority of users preferred using the front view first and then the side view. One user commented saying that the front view was "used for initial positioning of the grasp", and the side view was "used for detailed adjustments". Another user also commented "side view first to work out depth and grasp width, and then front view to confirm position". Some users even had a three step approach, one user commented saying "side view first, then adjust according to front view, then double check with side view".

Hypothesis - Revisited

$H_{2.1}$: using dual visual feedback in grasping virtual objects that change in size has no effect on a) grasp aperture and b) grasp displacement: **Rejected** as statistically significant results were found for the feedback method condition showing that dual visual feedback in grasping virtual objects that change in size has a significant effect on GAp, and on $GDisp$ in all axes (x, y and z).

6.4 Experiment 2 - Object Position

6.4.1 Experiment Design

A 2 × 3 × 3 × 3 within-subjects design is used, with two primary conditions: object position and object type (see Table 6.1 [page 102]). All 15 new participants took part in both conditions. Every permutation of position for both object types was randomly presented to participants to exclude potential learning effects. In total, each participant completed 27 (positions) × 2 (objects) = 54 trials and 810 grasps (54 trials × 15 participants). Each static grasp of every participant was recorded for 5 seconds (75 frames), leading to collecting 60750 raw data points (810 grasps × 75 frames).

6.4.2 Procedure

For this second experiment, participants were instructed to accurately locate and match their grasp aperture to the size and position of the virtual object in the shortest time possible on both feedback views. 27 different positions in all axes (x, y and z) are used (see Table 6.1 [page 102]), covering a working range of 400mm from participants (see Figure 6.1a [page 101]). The object sizes chosen for this experiment were the two sizes that had the lowest mean difference between

GAp and object size in Study 1 (Chapter 5) (80mm for cubes and 70mm for spheres) and were unchanged throughout the experiment.
Before interaction, an object (cube or sphere) appeared to participants on the feedback monitor, each object had 27 different positions. A countdown of 5 seconds followed by an auditory cue was used as an indicator for participants to start grasping the object
During the experiment, all participants were instructed to verbally inform the test coordinator that they are satisfied with the grasp they have performed on both feedback views (front and side), and maintain the grasp for 5 seconds while the measurements are stored.

6.4.3 Results

The object position that was used in Experiment 1 (Centre) was changed in the x, y and z axes (see Table 6.1 [page 102]). In order to directly compare the two experiments in this study, this section will only report on and analyse results of the z plane that was used in Experiment 1 (1600mm), and changes in object positions were compared as whole sets between the two feedback methods (single and dual view visual feedback) to test the influence of the proposed visual feedback method in this study on GAp and $GDisp$ given that object position changes.

6.4.3.1 Results - Grasp Aperture (GAp)

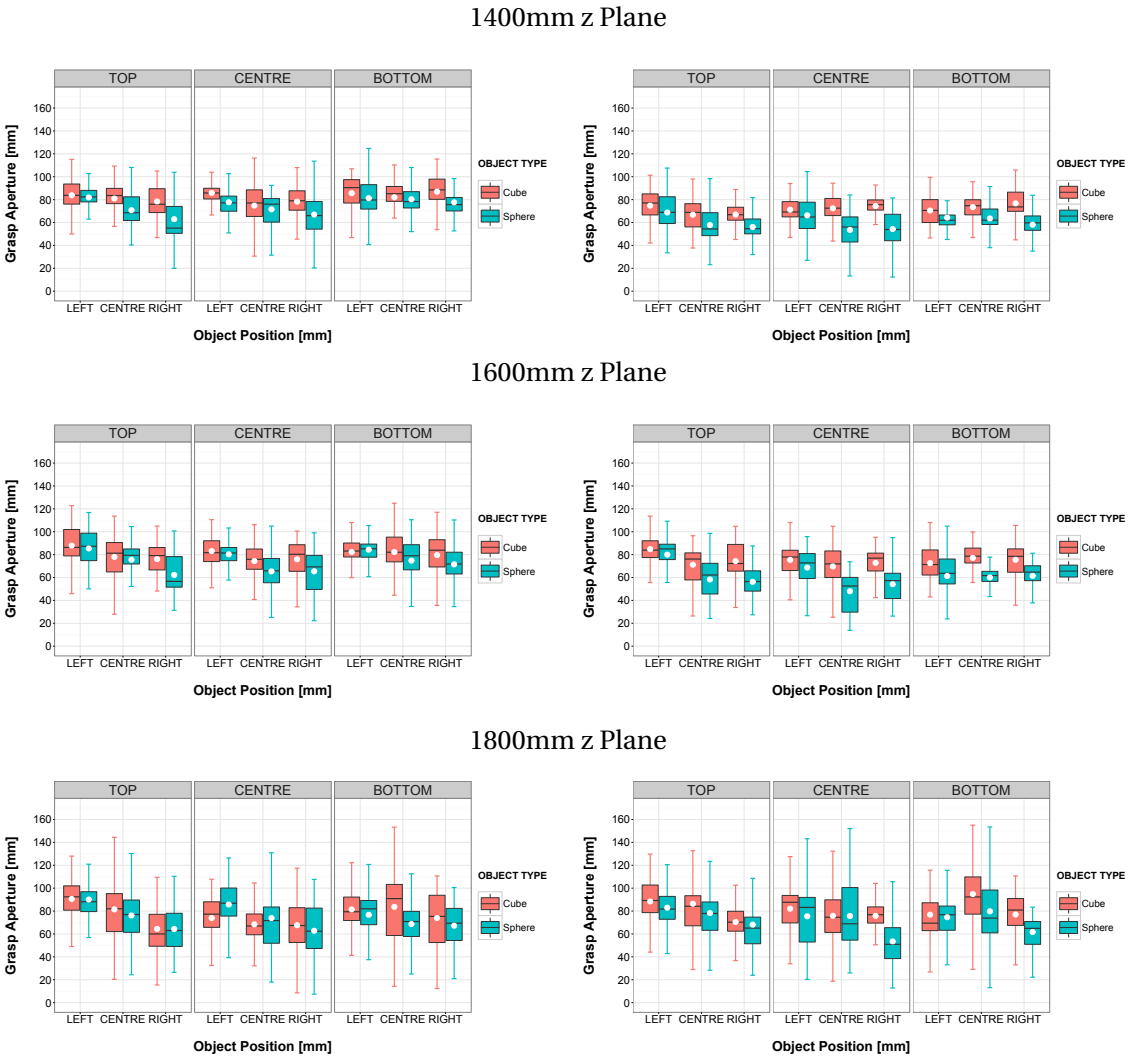
1400m Z plane
A statistically significant difference in GAp between the single and dual view visual feedback conditions in different positions was found for cubes ($\chi^2(1) = 2900$, $p < 0.01$) and spheres ($\chi^2(1) = 3993$, $p < 0.01$)
1600m Z plane
A statistically significant difference in GAp between the single and dual view visual feedback conditions in different positions was found for cubes ($\chi^2(1) = 648$, $p < 0.01$) and spheres ($\chi^2(1) = 2508$, $p < 0.01$)
1800m Z plane
A statistically significant difference in GAp between the single and dual view visual feedback conditions in different positions was found for cubes ($\chi^2(1) = 144$, $p < 0.01$) and spheres ($\chi^2(1) = 113$, $p < 0.01$)

6.4.3.2 Analysis - Grasp Aperture (GAp)

Users consistently underestimated object size in this experiment under the dual view visual feedback condition in the majority of position, where underestimation of object size occurred in 69 out of the 81 trials under test (see Table 6.3 [page 115]). Interestingly this contradicts findings for single view visual feedback where users showed a consistent overestimation of object size and potentially shows that dual view visual feedback leads users to penetrate the bounds of virtual objects presented to achieve higher accuracy in size estimation, as users are able to clearly visualise their full hand along the z axis.

Table 6.3: Descriptive Statistics of GAp (Mean ± SD. Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*)

1400mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	Single	83.82 ± 15.30*	80.98 ± 13.84*	78.40 ± 13.15*
		Dual	74.77 ± 11.73*	66.81 ± 12.29*	66.98 ± 12.41*
	Centre	Single	85.92 ± 10.98*	74.77 ± 19.50*	78.24 ± 13.95*
		Dual	71.19 ± 9.77*	72.39 ± 11.02*	74.37 ± 10.37*
	Bottom	Single	85.58 ± 13.90*	82.06 ± 20.11*	86.95 ± 14.88*
		Dual	70.48 ± 11.46*	73.45 ± 10.70*	76.71 ± 11.85*
Spheres (Constant Size - 70mm)	Top	Single	81.92 ± 9.80*	70.80 ± 14.69*	62.89 ± 17.82*
		Dual	68.75 ± 15.64*	57.76 ± 16.98*	56.13 ± 8.96*
	Centre	Single	77.71 ± 13.15*	71.51 ± 11.95*	66.96 ± 19.53*
		Dual	66.26 ± 15.22*	53.55 ± 14.82*	54.25 ± 13.80*
	Bottom	Single	81.27 ± 14.58*	80.30 ± 10.82*	77.64 ± 12.96*
		Dual	64.28 ± 11.51*	63.55 ± 12.55*	58.09 ± 11.77*
1600mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	87.91 ± 17.41*	77.96 ± 17.00*	76.04 ± 14.74
		Dual	84.80 ± 12.73*	71.21 ± 13.59*	74.31 ± 16.21
	Centre	Single	83.13 ± 12.39*	74.36 ± 16.84*	75.92 ± 15.10*
		Dual	75.37 ± 11.39*	69.69 ± 16.80*	72.80 ± 13.00*
	Bottom	Single	82.28 ± 13.79*	82.31 ± 15.33*	79.56 ± 17.45*
		Dual	72.59 ± 12.50*	77.10 ± 13.60*	75.53 ± 15.32*
Spheres (Constant Size - 70mm)	Top	Single	85.36 ± 15.59*	75.54 ± 16.69*	62.27 ± 15.33*
		Dual	79.97 ± 17.52*	58.29 ± 17.53*	56.24 ± 13.35*
	Centre	Single	80.25 ± 10.23*	65.34 ± 15.76*	65.44 ± 17.66*
		Dual	68.70 ± 16.40*	48.00 ± 18.59*	54.14 ± 14.85*
	Bottom	Single	84.33 ± 14.58*	74.77 ± 18.68*	71.50 ± 16.72*
		Dual	61.30 ± 19.02*	60.05 ± 12.65*	61.48 ± 12.85*
1800mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	90.60 ± 17.35	81.46 ± 26.56*	64.46 ± 18.39*
		Dual	88.38 ± 17.76	86.38 ± 29.56*	70.46 ± 16.29*
	Centre	Single	73.93 ± 19.51*	68.41 ± 19.09*	67.62 ± 20.60*
		Dual	82.06 ± 18.72*	75.90 ± 25.15*	75.79 ± 19.50*
	Bottom	Single	81.46 ± 20.46*	83.71 ± 26.17*	73.88 ± 22.48
		Dual	76.84 ± 23.35*	94.86 ± 26.87*	76.95 ± 18.17
Spheres (Constant Size - 70mm)	Top	Single	90.06 ± 16.27*	76.13 ± 20.92	64.33 ± 24.46
		Dual	82.96 ± 17.64*	78.25 ± 23.15	68.11 ± 29.20
	Centre	Single	85.72 ± 20.27*	73.91 ± 31.14	62.81 ± 20.80*
		Dual	75.41 ± 24.06*	75.67 ± 30.98	53.51 ± 21.00*
	Bottom	Single	76.75 ± 20.32*	68.62 ± 21.99*	67.23 ± 16.18*
		Dual	74.60 ± 16.41*	79.80 ± 30.21*	61.72 ± 15.28*



(a) Single View

(b) Dual View

Figure 6.5: GAp for different object positions in the three z planes in this experiment (1400mm, 1600mm and 1800mm). 6.5a: Single view visual feedback. 6.5b: Dual view visual feedback. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

Figure 6.6: Users showed higher accuracy in terms of matching GAp to objects size in the closest 1800mm z plane. This is potentially attributed to the placement of the object in relation to the body of the user, where the body in this particular plane acts as a spatial cue for the virtual object presented

Users again showed wide working ranges of mean GAp across all z planes for cubes (from 66.81 ± 12.29 to 94.86 ± 26.87) and spheres (from 48.00 ± 18.59 to 82.96 ± 17.64), despite the constant sizes of the two objects in this experiment (see Figure 6.5b [page 116]). This shows that freehand grasping is still prone to high variation in GAp by users even with the addition of a second view for visual feedback, as this behaviour was also present for single view visual feedback (for cubes (from 64.46 ± 18.39 to 90.60 ± 17.35) and spheres (from 62.27 ± 15.33 to 90.06 ± 16.27)) (see Figure 6.5a [page 116]), and can be attributed to the lack of tactile feedback in exocentric AR environments.

Users showed statistically significant differences in GAp between different z planes under the dual view visual feedback condition for cubes ($\chi^2(2) = 1293$, $p < 0.01$) and spheres ($\chi^2(2) = 1621$, $p < 0.01$), and presented the highest accuracy in terms of matching their GAp to object sizes in the 1800mm z plane for cubes (mean overestimation of 0.85mm ± 23.30) and spheres (mean overestimation of 2.22mm ± 25.40) in the dual view visual feedback condition (see Table 6.3 [page 115]). This was surprising as users consistently showed higher accuracy in the 1600mm z plane for single view visual feedback in Chapter 5 where this preference was attributed to the convenience of the 1600mm z plane as it requires the least amount of arm flexion and extension (see Figure 6.5a [page 116]). For dual view visual feedback, this preference for the 1800mm z plane can be attributed to the test design, where objects in this particular plane were placed close to, or in some cases on the head of the user (this was largely dependent on the amount of movement the user made away from the instructed standing position in this study using their upper body (e.g. bend forward or backward), and not their full body using their legs as users were instructed not move in this study and were obstructed from doing so using a mark and a physical box on the floor) (see Figure 6.6 [page 117]), this potentially presented users with an additional cue regarding the position of the virtual object, that cue or reference being the body of the user. This knowledge about the position of the object in relation to the body of the user is acquired by users via the second view for visual feedback before they start the grasping interaction, this accordingly limits the time users spend in locating the object and arm flexion and potentially leads to a higher accuracy in size estimation (see Figure 6.5b [page 116]).

As shown in Figure 6.7f [page 119], users showed the highest accuracy in matching GAp to object size in the Bottom positions alongside the y axis for cubes with a mean underestimation of -2.83mm $\pm$ 18.11, and in the Top positions for spheres with a mean underestimation of -2.61mm $\pm$ 21.22. Preference for top positions is in alignment with that of users under the single view visual feedback where users showed the highest accuracy in Top and Centre positions along the y axis for both objects. Users potentially preferred Bottom positions for cubes in the dual view visual feedback conditions due to the restricting nature of this position that was more physically demanding in terms of arm reach and body position in comparison to top and centre positions, thus the inconvenience of bottom positions, in this case, limited the amount of error in size estimation by users.

Alongside the x axis (see Figure 6.7a [page 119]), users showed the highest accuracy in matching GAp to object size in the Left positions for cubes (mean underestimation of -2.61mm $\pm$ 16.12) and spheres (mean overestimation of 1.36mm $\pm$ 18.65). This differs from single view visual feedback, where users consistently showed higher accuracy in matching GAp to object size in Centre and Right positions, and the least accuracy was consistently found in Left positions (see Table 6.3 [page 115]). This is potentially due to the experiment design in this study where the window showing the side view to users was placed on the right hand of users, and given that all users are right handed in this study, reaching and grasping virtual objects placed in the Right positions was problematic due to the position of the arm that intersects or blocks the view of users from the side view visual feedback on the main feedback monitor, this was particularly prominent in the closest 1800mm z plane to users (see Figure 6.8 [page 120]). This problem of users obstructing the view of their interaction using their arm highlights a limitation in dual view visual feedback as it can potentially hinder usability and grasping accuracy. One potential solution to this problem that can be addressed in future work would be to change the position of the side view window depending on the position of the virtual object in the scene, meaning that that the side view window on the feedback monitor would be placed on the left if the virtual object is on the right hand side of users, and on the right of the feedback monitor if the virtual object is on the left hand side of users.

In terms of accuracy, users performed better in matching GAp to object size in the single view visual feedback condition than the dual view condition in the majority of individual positions across the three z planes for both objects (see Table 6.3 [page 115]), where dual view visual feedback outperformed single view visual feedback in only 17 out of the 54 trials in this experiment. However, users under the dual view visual feedback condition showed less variation in GAp for the majority of positions in this experiment (31 out of 54 trials in this experiment) (see Table 6.3 [page 115]), thus users were more confident in matching object size using their GAp when provided with dual view visual feedback but not more accurate. Participants performed better in matching their GAp to object sizes using single view visual feedback (see Figures 6.5a [page 116] and 6.5b [page 116]), and this can be attributed to the fact that virtual objects changed position in this experiment, and as participants had no prior knowledge about the positions of the virtual

(a) GAp - X axis (b) Time - X axis (c) $GDisp_x$ - X axis

(d) $GDisp_y$ - X axis (e) $GDisp_z$ - X axis

(f) GAp - Y axis (g) Time - Y axis (h) $GDisp_x$ - Y axis

(i) $GDisp_y$ - Y axis (j) $GDisp_z$ - Y axis

Figure 6.7: Optimal interaction regions for users across all z planes in terms of all the measurement used in this work to assess grasp accuracy. X axis refers to Left, Centre and Right positions. Y axis refers to Top, Centre and Bottom positions. Values presented are Means ± SD of each corresponding measurement. Most accurate/quickest region is marked with a star

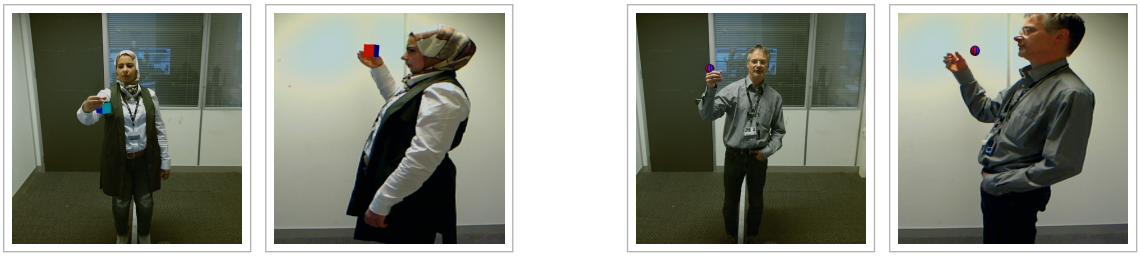


Figure 6.8: Users consistently showed the lowest accuracy in terms of matching GAp to objects size in Right positions. This is potentially attributed to the position of the side view window on the feedback monitor that is on the right hand side of users, and this lead users to move their upper body or just their heads in order to be able to correct their GAp

objects that are presented in this experiment, accurately locating virtual objects in 3D space using their gmp was prioritised over accurately match their GAp to object size. Even though this behaviour was present in single view visual feedback, presenting second view visual feedback to participants made this behaviour more prominent.

6.4.3.3 Results - Completion Time

1400m Z plane

A statistically significant difference in GAp between the single and dual view visual feedback conditions in different positions was found for cubes ($\chi^2(1) = 3830$, $p < 0.01$) and spheres ($\chi^2(1) = 4752$, $p < 0.01$)

1600m Z plane

A statistically significant difference in GAp between the single and dual view visual feedback conditions in different positions was found for cubes ($\chi^2(1) = 5778$, $p < 0.01$) and spheres ($\chi^2(1) = 6212$, $p < 0.01$).

1800m Z plane

A statistically significant difference in GAp between the single and dual view visual feedback conditions in different positions was found for cubes ($\chi^2(1) = 5107$, $p < 0.01$) and spheres ($\chi^2(1) = 6945$, $p < 0.01$)

6.4.3.4 Analysis - Completion Time

Mean completion time ranged across all z planes and positions in the dual view feedback condition from 8.47s ± 3.08 to 20.00s ± 19.83 for cubes, and from 7.53s ± 2.92 to 18.27s ± 11.86 for spheres (see Table 6.4 [page 121]). These ranges are significantly higher than the ones shown by users under the single view visual feedback condition (from 5.00s ± 3.74 to 11.40s ± 13.80 for cubes, and from 4.06s ± 2.05 to 7.53s ± 4.15 for spheres). This was expected as providing users with an additional view of their interaction in this study naturally presents a higher cognitive load and leads users to spend more time in adjusting their reaching and grasping parameters. Users showed statistically significant differences in task completion time between different z

Table 6.4: Descriptive Statistics of Task Completion Time (Mean ± SD). Significant differences (p < 0.01) between single and dual view visual feedback methods are marked with (*)

1400mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	Single	7.20 ± 5.67*	5.46 ± 2.25*	6.60 ± 3.65*
		Dual	14.40 ± 8.64*	10.40 ± 6.79*	11.66 ± 4.47*
	Centre	Single	5.93 ± 3.34*	5.27 ± 2.18*	9.20 ± 6.15*
		Dual	13.40 ± 7.62*	8.60 ± 4.46*	11.00 ± 5.38*
	Bottom	Single	6.67 ± 3.61*	5.66 ± 2.09*	9.20 ± 5.78*
		Dual	10.33 ± 3.96*	11.46 ± 9.19*	11.93 ± 9.38*
Spheres (Constant Size - 70mm)	Top	Single	6.40 ± 3.24*	4.46 ± 2.03*	7.53 ± 4.15*
		Dual	10.00 ± 3.31*	9.13 ± 4.05*	7.53 ± 2.92*
	Centre	Single	5.20 ± 2.54*	5.60 ± 4.50*	5.73 ± 2.44*
		Dual	10.40 ± 7.47*	7.73 ± 3.51*	10.60 ± 5.42*
	Bottom	Single	5.73 ± 2.98*	5.20 ± 2.29*	6.87 ± 3.58*
		Dual	9.27 ± 4.25*	10.67 ± 5.99*	10.67 ± 6.86*
1600mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	6.73 ± 5.79*	5.93 ± 4.11*	8.46 ± 8.46*
		Dual	11.80 ± 6.21*	13.93 ± 19.66*	12.60 ± 6.11*
	Centre	Single	6.93 ± 7.26*	5.33 ± 3.34*	6.00 ± 4.73*
		Dual	11.60 ± 4.56*	10.00 ± 5.58*	10.33 ± 5.17*
	Bottom	Single	6.00 ± 3.88*	5.00 ± 3.74*	6.27 ± 2.54*
		Dual	11.73 ± 5.92*	8.47 ± 3.08*	12.40 ± 7.99*
Spheres (Constant Size - 70mm)	Top	Single	5.80 ± 3.37*	5.13 ± 3.33*	4.60 ± 1.36*
		Dual	10.53 ± 5.79*	9.13 ± 5.70*	14.33 ± 8.96*
	Centre	Single	4.53 ± 2.68*	4.46 ± 1.96*	5.00 ± 2.39*
		Dual	11.00 ± 5.82*	10.00 ± 6.41*	8.66 ± 5.22*
	Bottom	Single	5.80 ± 2.20*	4.53 ± 2.06*	6.40 ± 2.89*
		Dual	9.80 ± 5.33*	7.60 ± 2.82*	12.60 ± 7.70*
1800mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	11.40 ± 13.80*	5.07 ± 2.77*	6.13 ± 3.44*
		Dual	20.00 ± 19.83*	12.40 ± 7.29*	19.20 ± 11.23*
	Centre	Single	7.93 ± 8.12*	6.13 ± 4.17*	6.07 ± 3.91*
		Dual	11.47 ± 6.56*	8.73 ± 4.77*	11.33 ± 5.16*
	Bottom	Single	6.60 ± 3.85*	5.73 ± 3.22*	6.13 ± 3.33*
		Dual	9.20 ± 3.58*	10.80 ± 6.33*	10.20 ± 4.79*
Spheres (Constant Size - 70mm)	Top	Single	6.00 ± 4.17*	5.93 ± 4.78*	5.67 ± 2.41*
		Dual	14.67 ± 8.38*	9.40 ± 4.56*	18.27 ± 11.86*
	Centre	Single	4.73 ± 2.86*	4.06 ± 2.05*	5.80 ± 3.29*
		Dual	11.87 ± 7.38*	8.47 ± 4.65*	12.07 ± 7.10*
	Bottom	Single	5.47 ± 2.90*	4.20 ± 2.04*	5.06 ± 2.24*
		Dual	14.40 ± 10.10*	9.47 ± 4.70*	8.40 ± 4.93*

planes under the dual view visual feedback condition for cubes ($\chi^2(2) = 101$, $p < 0.01$) and spheres ($\chi^2(2) = 492$, $p < 0.01$), and users showed the shortest mean completion times in the 1600mm z plane in grasping cubes (11.43s $\pm$ 8.63) and in the 1400mm z plane in grasping spheres (9.56s $\pm$ 5.23) (see Table 6.4 [page 121]). Interestingly, users showed the highest accuracy in terms of matching GAp to object size in the 1800mm z plane where users showed the highest mean completion times for both objects. This shows that mean completion time is not necessarily a reliable measure for accuracy in freehand grasping. However, completion times are useful in determining the usability of dual view visual feedback as it significantly increased mean completion times in all z planes in comparison to single view visual feedback.

As shown in Figure 6.7g [page 119], users showed shortest mean completion times across all z planes in the Centre positions for cubes (10.72s $\pm$ 5.73) and spheres (10.09s $\pm$ 6.18). Central interaction regions are usually preferred by users in freehand grasping as they are easily accessible using the right dominant hand and reach to grasp movements. However, the highest accuracy in size estimation using GAp was found in Bottom (for cubes) and Top (spheres) positions.

Similarly alongside the x axis (see Figure 6.7b [page 119]), users also showed shortest mean completion times across all z planes in the Centre positions for cubes (10.53s $\pm$ 8.95) and spheres (9.07s $\pm$ 4.93). This is in alignment with findings for single view visual feedback, and even though users showed the highest accuracy in matching object size in the Left positions, this again highlights the convenience of Centre position for freehand grasping interactions.

6.4.3.5 Results - Grasp Displacement ($GDisp$)

1400m Z plane

A statistically significant difference between the single and dual view visual feedback conditions in different positions was found in $GDisp_{\text{x}}$ (for cubes ($\chi^2(1) = 253$, $p < 0.01$) and spheres ($\chi^2(1) = 285$, $p < 0.01$)), $GDisp_{\text{y}}$ (for cubes ($\chi^2(1) = 721$, $p < 0.01$) and spheres ($\chi^2(1) = 235$, $p < 0.01$)) and $GDisp_{\text{z}}$ (for cubes ($\chi^2(1) = 3587$, $p < 0.01$) and spheres ($\chi^2(1) = 2680$, $p < 0.01$)).

1600m Z plane

A statistically significant difference between the single and dual view visual feedback conditions in different positions was found in $GDisp_{\text{x}}$ (for cubes ($\chi^2(1) = 210$, $p < 0.01$) and spheres ($\chi^2(1) = 23$, $p < 0.01$)), $GDisp_{\text{y}}$ (for cubes ($\chi^2(1) = 3026$, $p < 0.01$) and spheres ($\chi^2(1) = 1349$, $p < 0.01$)) and $GDisp_{\text{z}}$ (for cubes ($\chi^2(1) = 2298$, $p < 0.01$) and spheres ($\chi^2(1) = 1990$, $p < 0.01$)).

1800m Z plane

A statistically significant difference between the single and dual view visual feedback conditions in different positions was found in $GDisp_{\text{x}}$ (for cubes ($\chi^2(1) = 173$, $p < 0.01$) but not spheres ($\chi^2(1) = 6.55$, $p > 0.01$)), $GDisp_{\text{y}}$ (for cubes ($\chi^2(1) = 2172$, $p < 0.01$) and spheres ($\chi^2(1) = 822$, $p < 0.01$)) and $GDisp_{\text{z}}$ (for cubes ($\chi^2(1) = 2066$, $p < 0.01$) and spheres ($\chi^2(1) = 1766$, $p < 0.01$)).

6.4.3.6 Analysis - Grasp Displacement ($GDisp$)

***GDisp*$_{\text{x}}$**

Table 6.5: Descriptive Statistics of $GDisp_x$ (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*)

1400mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	Single	45.82 ± 25.96*	33.00 ± 20.56*	10.53 ± 22.32
		Dual	34.24 ± 23.74*	20.11 ± 18.44*	10.76 ± 18.47
	Centre	Single	30.68 ± 17.96*	28.28 ± 22.35*	20.11 ± 26.80
		Dual	25.66 ± 15.74*	24.55 ± 16.97*	18.80 ± 13.17
	Bottom	Single	31.69 ± 22.53*	27.28 ± 24.10*	17.44 ± 21.94
		Dual	27.96 ± 17.11*	22.83 ± 21.85*	22.12 ± 18.20*
Spheres (Constant Size - 70mm)	Top	Single	45.30 ± 27.97*	23.56 ± 20.07*	1.03 ± 23.67*
		Dual	39.41 ± 36.11*	19.23 ± 16.33*	18.68 ± 12.37*
	Centre	Single	28.91 ± 17.67*	21.64 ± 17.28*	13.07 ± 27.16*
		Dual	40.38 ± 44.75*	34.23 ± 30.91*	24.04 ± 37.30*
	Bottom	Single	42.71 ± 45.76	20.22 ± 24.01*	12.99 ± 18.57*
		Dual	42.21 ± 44.34	32.26 ± 32.43*	22.13 ± 17.84*
1600mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	52.28 ± 26.93*	34.77 ± 23.36*	5.94 ± 30.31
		Dual	42.34 ± 19.10*	21.90 ± 21.69*	6.69 ± 14.94
	Centre	Single	37.44 ± 18.57*	34.81 ± 19.01*	13.49 ± 25.95*
		Dual	41.90 ± 21.60*	29.36 ± 18.39*	13.05 ± 12.29*
	Bottom	Single	38.82 ± 22.70	32.75 ± 25.12*	23.18 ± 14.83*
		Dual	36.90 ± 19.10	25.36 ± 24.20*	18.48 ± 18.58*
Spheres (Constant Size - 70mm)	Top	Single	47.93 ± 33.40*	21.88 ± 17.66*	-2.96 ± 24.53*
		Dual	37.88 ± 32.83*	33.62 ± 21.14*	7.54 ± 17.73*
	Centre	Single	40.44 ± 25.24*	25.64 ± 17.02*	7.32 ± 25.95*
		Dual	46.34 ± 32.12*	19.09 ± 19.65*	14.96 ± 15.03*
	Bottom	Single	38.71 ± 27.93	31.17 ± 26.45*	8.75 ± 14.61*
		Dual	43.48 ± 45.99	30.12 ± 33.57*	14.90 ± 17.88*
1800mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	58.82 ± 31.81*	27.29 ± 29.03*	-5.47 ± 28.86
		Dual	44.77 ± 38.86*	41.11 ± 34.39*	-2.23 ± 16.59
	Centre	Single	48.19 ± 29.52*	33.71 ± 16.12*	2.24 ± 28.51
		Dual	43.26 ± 32.77*	25.51 ± 18.03*	2.07 ± 33.73
	Bottom	Single	50.45 ± 24.43	37.12 ± 22.81*	17.63 ± 26.03*
		Dual	54.10 ± 40.07	31.67 ± 28.34*	11.84 ± 20.42*
Spheres (Constant Size - 70mm)	Top	Single	78.03 ± 41.89*	27.33 ± 32.94*	-7.56 ± 32.83*
		Dual	42.12 ± 27.90*	39.42 ± 34.04*	0.08 ± 22.64*
	Centre	Single	57.21 ± 43.80*	27.78 ± 19.05	4.13 ± 18.35*
		Dual	48.84 ± 39.03*	28.74 ± 20.19	16.69 ± 33.57*
	Bottom	Single	56.63 ± 36.94	33.16 ± 34.30*	6.69 ± 28.07*
		Dual	60.90 ± 50.47	42.22 ± 43.57	7.51 ± 22.30*

$GDisp_x$ showed the highest displacement in this experiment for both objects under the dual view visual feedback condition, unlike single view visual feedback where it was the second highest behind $GDisp_z$. Users showed a mean $GDisp_x$ of 25.75mm ± 26.97 in grasping cubes, and a mean of 29.89mm ± 34.68 for spheres across all positions and z planes.

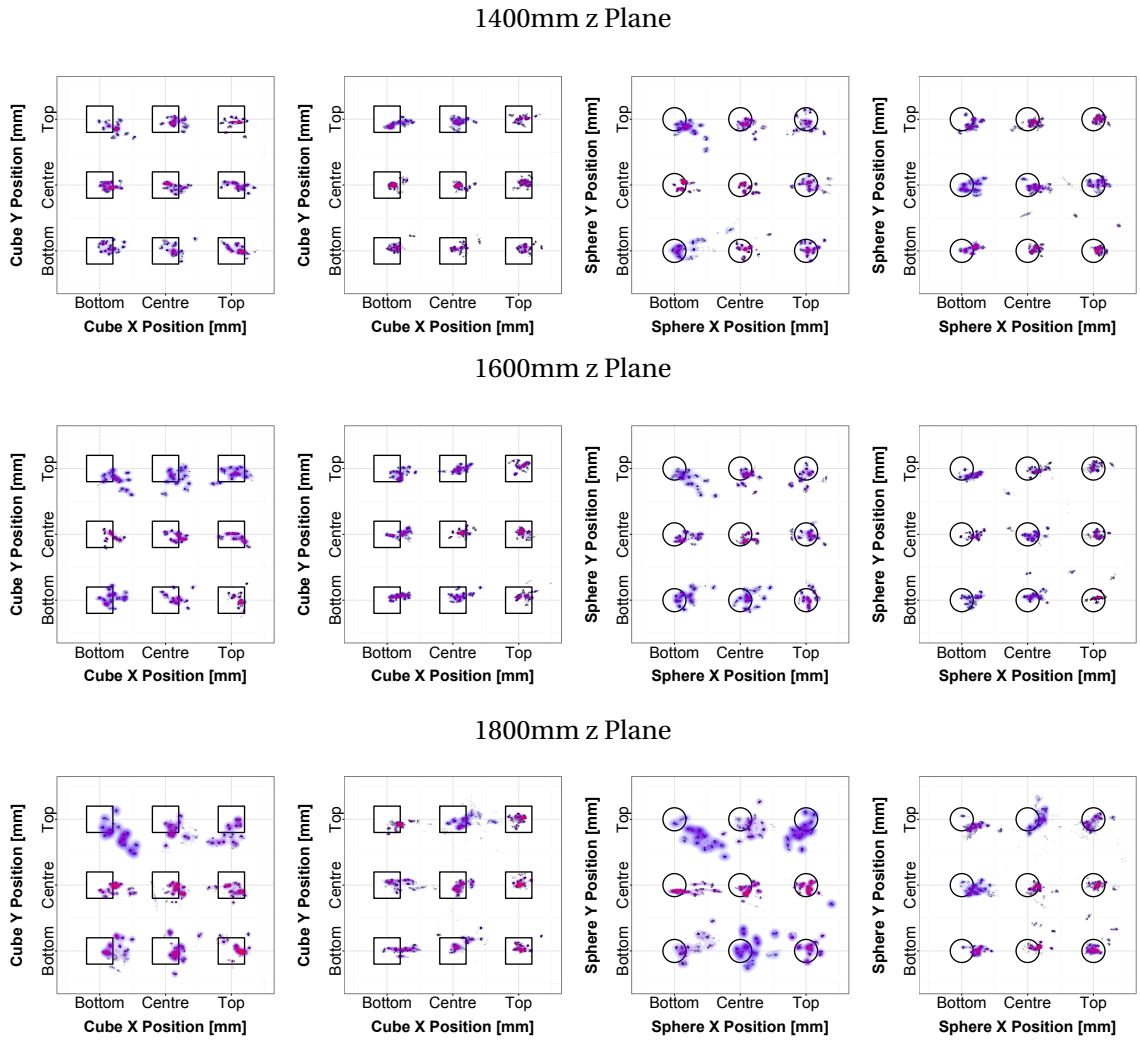
Users showed lower mean $GDisp_x$ in the dual view visual feedback condition in the majority of positions, where users showed lower $GDisp_x$ using dual view than single view visual feedback in 28 out of the 54 trials and also reduced variation in gmp placement along the x axis in 28 out of the 54 trials in this study (see Table 6.5 [page 123]). Adding a second view for visual feedback also reduced the range of $GDisp_x$ for cubes (ranged from 2.07mm ± 33.73 to 54.10mm ± 40.07) and spheres (ranged from 0.08mm ± 22.64 to 60.90 ± 50.47) in comparison to single view visual feedback where mean $GDisp_x$ ranged from -5.46mm ± 28.86 to 58.82 ± 31.81 for cubes and from -7.55mm ± 32.83 to 78.03 ± 32.83 for spheres (see Table 6.5 [page 123]).

As shown by the clusters in Figures 6.9b [page 125] and 6.9d [page 125], ranges of mean $GDisp_x$ were less spread for both objects using dual view visual feedback than single view visual feedback (see Figures 6.9a [page 125] and 6.9c [page 125]). Moreover, lower SD values show that less variability by participants in spatial placement of gmp in the x axis was also present while using dual view visual feedback (see Table 6.5 [page 123]). This shows that using dual view visual feedback shifted gmp of users in the x axis more towards the omp, and significantly reduced mean $GDisp_x$.

Users showed statistically significant differences in $GDisp_x$ between different z planes under the dual view visual feedback condition for cubes ($\chi^2(2) = 100$, $p < 0.01$) and spheres ($\chi^2(2) = 95$, $p < 0.01$), and presented the lowest $GDisp_x$ in the 1400mm z plane for cubes (23.00 ± 19.41) and in the 1600mm z plane for spheres (27.55 ± 30.89) (see Table 6.5 [page 123]). Preference for the furthest z plane (1400mm) is attributed to the position of the plane being at the extremity of the interaction region in this experiment, this limited the amount of errors users can make due to the limited possible arm reach or hand movements in this plane. Whereas the preference for the middle z plane (1600mm) is attributed to the spatial convenience of this plane as it requires less flexion and extension of the arm in comparison to the other two planes, thus it feels most natural to users.

As shown in Figure 6.7h [page 119], lowest mean $GDisp_x$ was shown by users in the Top positions alongside the y axis across all z planes for both cubes (24.41 ± 29.09) and spheres (26.44 ± 29.70). Users consistently preferred Top and Centre positions in gmp placement along the x axis as the two positions require the least amount of body movement.

Alongside the x axis (see Figure 6.7c [page 119]), users showed the lowest $GDisp_x$ in Right positions across all z planes for cubes (11.29 ± 20.84) and spheres (14.06 ± 24.34), this preference for gmp placement along the x axis in right positions is in alignment with findings for single view visual feedback, and is potentially attributed to the right-handedness of all users in this study. Interestingly, users showed the least accuracy in terms of matching GAp to object sizes in Right positions in this experiment, this indicates that gmp placement and GAp are impacted by differ-



(a) Single View - Cubes (b) Dual View - Cubes (c) Single View - Spheres (d) Dual View - Spheres

Figure 6.9: gmp placement in the x and y axes for cubes and spheres of all participants in 3 z planes (1400mm - 1600mm - 1800mm). 6.9a and 6.9c: Single view visual feedback. 6.9b and 6.9d: Dual view visual feedback. Density heat maps indicate gmp placement across participants (red indicates higher density)

ent aspects of the interaction design. This is evident in this study as GAp is highly impacted by the feedback method used, as users showed the highest accuracy in positions (Left) that allowed users to clearly see additional feedback that is the second view on the feedback monitor without the arm blocking their view in any way. In contrast for gmp placement, this is highly impacted by the convenience of the position of virtual objects in relation to the position of users and their dominant hand.

***GDisp*$_y$**

Similar to single view visual feedback, $GDisp_y$ showed the lowest displacement in this experiment for both objects under the dual view visual feedback condition. Users showed a mean $GDisp_y$ of 1.91mm ± 16.01 in grasping cubes, and a mean of -0.67mm ± 20.99 for spheres across all positions and z planes.

Users showed lower mean $GDisp_y$ in the dual view visual feedback condition in the majority of positions, where users showed lower $GDisp_y$ using dual view than single view visual feedback in 37 out of the 54 trials and also reduced variation in gmp placement along the x axis in 44 out of the 54 trials in this study (see Table 6.6 [page 127]). In addition, dual view visual feedback also reduced the number of negative mean $GDisp_y$ observations from 34 (in single view visual feedback) to 24, thus users shifted their gmp placement along the y axis more towards the omp.

As shown by the clusters in Figures 6.9b [page 125] and 6.9d [page 125], adding a second view for visual feedback also reduced the range of $GDisp_y$ for cubes (ranged from -19.85mm ± 11.57 to 22.86mm ± 18.38) and spheres (ranged from -23.97mm ± 15.70 to 20.84 ± 26.14) in comparison to single view visual feedback where mean $GDisp_y$ ranged from -47.82mm ± 53.81 to 11.59mm ± 13.15 for cubes and from -54.92mm ± 30.26 to 13.19mm ± 32.93 for spheres (see Figures 6.9b [page 125] and 6.9d [page 125]).

Users showed statistically significant differences in $GDisp_y$ between different z planes under the dual view visual feedback condition for cubes ($\chi^2(2) = 689$, $p < 0.01$) and spheres ($\chi^2(2) = 182$, $p < 0.01$), and showed the lowest mean $GDisp_y$ in the 1800mm z plane for cubes (0.49mm ± 20.45) potentially due to its spatial position at the extremity of the arm reach, and in the 1600mm z plane for spheres (0.29mm ± 18.81) due to its lower requirements in terms arm flexion and extension in comparison to the 1400mm and 1800mm z planes.

As shown in Figure 6.7i [page 119], lowest mean $GDisp_y$ under the dual view visual feedback condition was shown by users in the Centre positions alongside the y axis across all z planes for both cubes (0.52 ± 14.74) and spheres (-3.50 ± 14.29). Centre positions are usually preferred by users as they are less physically demanding in terms of arm reach.

Alongside the x axis (see Figure 6.7d [page 119]), users showed the lowest $GDisp_y$ in Centre positions across all z planes for cubes (3.17 ± 15.81) and spheres (0.45 ± 22.06), again as it is easily reachable by users in this experiment setting. Similar to $GDisp_x$, these preferred positions y users were not the positions that yielded the highest accuracy in terms of matching object size with GAp, thus again showing that gmp spatial accuracy is dependent on the convenience of positions for in relation to the user position, handedness and arm reach.

Table 6.6: Descriptive Statistics of $GDisp_{\mathrm{y}}$ (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*)

1400mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	Single	-27.76 ± 16.51*	-9.91 ± 12.09*	-12.06 ± 15.73*
		Dual	-13.30 ± 9.13*	-8.75 ± 11.25*	0.46 ± 8.65*
	Centre	Single	-3.65 ± 9.07*	-6.45 ± 10.51*	-3.63 ± 9.67*
		Dual	0.47 ± 10.34*	-3.20 ± 7.41*	6.18 ± 9.65*
	Bottom	Single	11.59 ± 13.15	4.56 ± 17.62*	4.76 ± 11.45*
		Dual	10.69 ± 11.85	9.29 ± 10.61*	7.40 ± 9.26*
Spheres (Constant Size - 70mm)	Top	Single	-31.00 ± 21.58*	-19.24 ± 16.07*	-14.40 ± 21.35*
		Dual	-15.41 ± 14.66*	-12.35 ± 8.63*	-1.31 ± 9.08*
	Centre	Single	-2.54 ± 12.15*	-9.26 ± 11.96	1.25 ± 13.35*
		Dual	-3.70 ± 14.93*	-8.52 ± 8.25	6.41 ± 14.86*
	Bottom	Single	8.53 ± 23.16*	6.30 ± 17.19	6.08 ± 16.25*
		Dual	13.40 ± 27.25*	11.39 ± 26.96	6.64 ± 20.28*
1600mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	-34.62 ± 20.65*	-27.75 ± 22.98*	16.69 ± 15.20*
		Dual	-16.14 ± 12.72*	-1.43 ± 8.82*	7.01 ± 12.22*
	Centre	Single	-7.64 ± 13.67*	-7.60 ± 10.99*	-10.60 ± 11.50*
		Dual	1.83 ± 9.38*	5.73 ± 9.41*	4.15 ± 10.08*
	Bottom	Single	7.72 ± 17.47*	6.60 ± 17.16*	2.35 ± 20.22*
		Dual	10.81 ± 6.60*	13.59 ± 11.85*	12.38 ± 12.09*
Spheres (Constant Size - 70mm)	Top	Single	-35.15 ± 23.21*	-26.81 ± 19.04*	-26.98 ± 21.39*
		Dual	-23.97 ± 15.70*	-5.82 ± 13.44*	5.52 ± 9.97*
	Centre	Single	-9.82 ± 10.74*	-13.24 ± 10.97*	-2.96 ± 15.14*
		Dual	-1.92 ± 9.95*	-4.13 ± 10.24*	-0.89 ± 8.39*
	Bottom	Single	10.29 ± 20.18*	5.89 ± 21.00*	5.84 ± 19.84*
		Dual	7.54 ± 23.77*	16.11 ± 20.91*	10.17 ± 17.40*
1800mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	-47.63 ± 34.09*	-47.82 ± 53.81*	-44.18 ± 26.56*
		Dual	-19.85 ± 11.57*	-5.12 ± 16.80*	0.11 ± 11.72*
	Centre	Single	-12.69 ± 12.53*	-15.61 ± 16.18*	-14.98 ± 14.37*
		Dual	-16.52 ± 11.69*	-4.47 ± 12.55*	10.47 ± 25.04*
	Bottom	Single	6.31 ± 26.34	0.73 ± 32.73*	9.75 ± 21.34*
		Dual	5.97 ± 17.43	22.86 ± 18.38*	10.97 ± 12.29*
Spheres (Constant Size - 70mm)	Top	Single	-54.92 ± 30.26*	-23.32 ± 29.25*	-51.04 ± 62.83*
		Dual	-23.38 ± 14.00*	-3.93 ± 28.22*	-9.51 ± 15.97*
	Centre	Single	-15.08 ± 10.01*	-15.62 ± 15.78*	-7.89 ± 13.27*
		Dual	-12.95 ± 15.32*	-9.53 ± 10.31*	3.69 ± 20.00*
	Bottom	Single	13.19 ± 32.93*	9.43 ± 32.22*	9.11 ± 24.11
		Dual	7.62 ± 23.91*	20.84 ± 26.14	7.51 ± 22.30

These results show that dual view visual feedback outperforms single view visual feedback by reducing mean $GDisp_x$ and $GDisp_y$ across all positions which shifted placement of gmp towards the 0 origin of the x and y axis, and also by reducing the range of $GDisp_x$ and $GDisp_y$ with less deviation, meaning that participants were more consistent in their spatial gmp placement in the x and y axes using dual view visual feedback.

***GDisp*$_z$**

Similar to the Object Size Experiment, $GDisp_z$ presented the highest displacement out of all three axes with single view visual feedback in Study 1 (Chapter 5) for grasping virtual objects in different positions. Here mean $GDisp_z$ across all positions was 3.33mm (SD = 22.17) for cubes and 5.07mm (SD = 26.28) for spheres. This shows a significant improvement in spatial gmp placement in the z axis as reported $GDisp_z$ means for single view visual feedback were -58.75mm (SD = 94.90) for cubes, and -51.60mm (SD = 89.06) for spheres

$GDisp_z$ presented the highest displacement out of the three axis in the previous study (Chapter 5) and in the first experiment in this study (see Section 6.3 [page 104]). However, in this experiment, $GDisp_z$ presented the second highest displacement behind $GDisp_y$ using dual view visual feedback, where users showed a mean $GDisp_z$ of 3.09mm $\pm$ 27.93 for cubes and 6.74mm $\pm$ 38.03 for spheres across all z planes and positions in this experiment. These means are significantly lower than those observed under the single view visual feedback condition (-70.04mm $\pm$ 122.47 for cubes and -57.37mm $\pm$ 110.80 for spheres), and this shows a significant improvement in spatial gmp placement in the z axis due to dual view visual feedback.

Users showed significantly lower $GDisp_z$ means in the dual view visual feedback condition in the majority of positions, where users showed lower $GDisp_z$ using dual view than single view visual feedback in 37 out of the 54 trials, and also reduced $GDisp_z$ variation in 51 out of the 54 trials in this experiment (see Table 6.7 [page 129]), thus indicating that participants had less variability in their depth estimation across all positions when using dual view visual feedback.

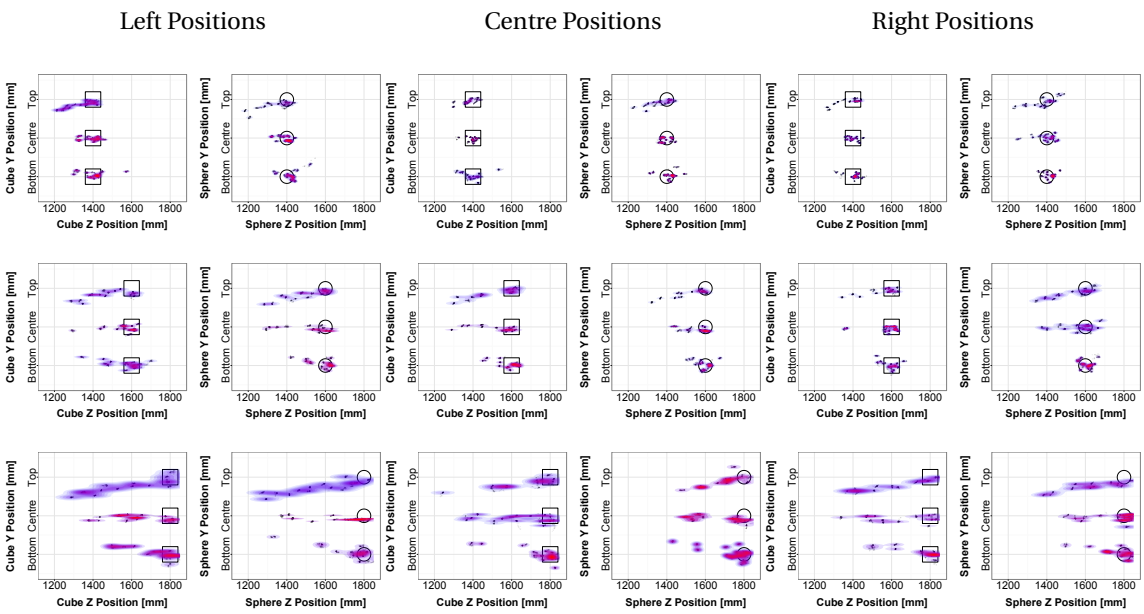
As shown by the clusters in Figure 6.10b [page 130], adding a second view for visual feedback shifted gmp placement along the z axis closer to the 0 origin, and significantly reduced the range of $GDisp_z$ for grasping both cubes (ranged from -41.73mm $\pm$ 29.17 to 34.88mm $\pm$ 15.62) and spheres (ranged from -40.36mm $\pm$ 18.98 to 47.52mm $\pm$ 47.09), in comparison to the ranges found for the single view visual feedback condition (from -220.12mm $\pm$ 181.69 to 5.94 $\pm$ 34.00 for cubes and from -233.25mm $\pm$ 174.06 to 20.55 $\pm$ 43.38 for spheres) (see Figure 6.10a [page 130]).

Interestingly dual view visual feedback also lead users to overestimate object position (positive $GDisp_z$) in the z axis in the majority of positions for both objects (34 out of 54 trials), in comparison to single view visual feedback where users underestimated object position in the majority of trials in this experiment (48 out of 54) (see Table 6.7 [page 129]). This overestimation shows that dual view visual feedback led participants to place their gmp in front of the omp.

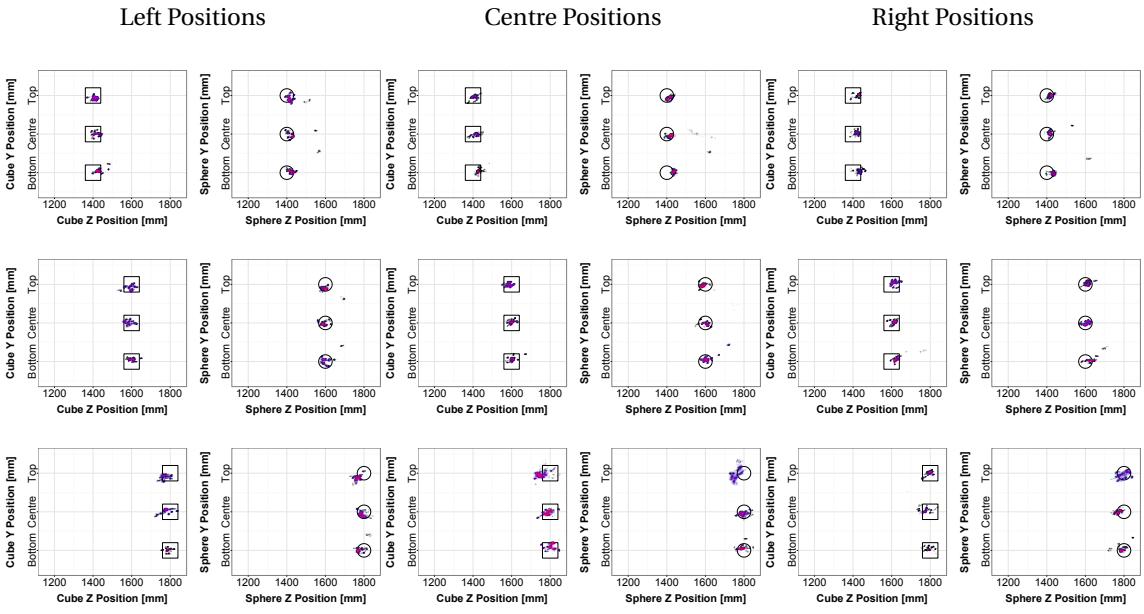
Users showed statistically significant differences in $GDisp_z$ between different z planes under the dual view visual feedback condition for cubes ($\chi^2(2) = 9804$, $p < 0.01$) and spheres ($\chi^2(2) = 11322$, $p < 0.01$), and showed the lowest mean $GDisp_z$ in the 1600mm z plane for cubes (3.34mm

Table 6.7: Descriptive Statistics of $GDisp_z$ (Mean ± SD). Significant differences ($p < 0.01$) between single and dual view visual feedback methods are marked with (*)

1400mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	Single	-65.70 ± 68.36*	-21.72 ± 38.82*	-20.03 ± 58.36*
		Dual	6.52 ± 13.41*	6.32 ± 14.30*	24.64 ± 13.75*
	Centre	Single	-12.26 ± 42.97*	-4.34 ± 25.51*	-0.98 ± 26.33*
		Dual	17.23 ± 15.77*	13.45 ± 16.75*	18.91 ± 10.71*
	Bottom	Single	0.73 ± 64.25*	-1.36 ± 50.85*	5.94 ± 34.00*
		Dual	32.77 ± 22.22*	30.21 ± 11.08*	34.88 ± 15.62*
Spheres (Constant Size - 70mm)	Top	Single	-60.44 ± 75.96*	-40.39 ± 59.44*	-19.48 ± 66.01*
		Dual	22.24 ± 25.81*	13.64 ± 12.41*	19.59 ± 10.93*
	Centre	Single	-12.11 ± 35.61*	-6.17 ± 26.68*	-1.77 ± 51.50*
		Dual	25.89 ± 35.34*	23.82 ± 37.12*	24.40 ± 29.62*
	Bottom	Single	20.55 ± 43.38	11.46 ± 43.57*	12.87 ± 28.38*
		Dual	31.73 ± 38.42	47.52 ± 47.09*	43.39 ± 47.80*
1600mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	-129.32 ± 105.30*	-98.24 ± 113.29*	-53.76 ± 85.90*
		Dual	-6.51 ± 21.42*	-13.36 ± 14.71*	15.58 ± 15.90*
	Centre	Single	-53.18 ± 86.73*	-85.24 ± 104.04*	-34.32 ± 83.12*
		Dual	-12.25 ± 20.30*	1.73 ± 14.59*	9.53 ± 13.58*
	Bottom	Single	-44.49 ± 88.57*	-22.88 ± 56.20*	-7.28 ± 38.87*
		Dual	3.15 ± 17.72*	6.62 ± 21.91*	25.56 ± 24.11*
Spheres (Constant Size - 70mm)	Top	Single	-106.78 ± 107.01*	-82.97 ± 87.37*	-72.44 ± 99.93*
		Dual	1.44 ± 27.22*	-12.92 ± 20.16*	12.80 ± 14.96*
	Centre	Single	-85.16 ± 113.66*	-49.60 ± 62.35*	-37.96 ± 87.15*
		Dual	-8.7 ± 18.18*	0.64 ± 18.34*	3.78 ± 16.26*
	Bottom	Single	-26.07 ± 62.43*	-8.20 ± 40.61*	4.77 ± 30.20*
		Dual	5.85 ± 27.76*	13.77 ± 34.81*	28.98 ± 26.04*
1800mm Z Plane					
Object Type	Position (y)	View	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	Single	-220.12 ± 181.69*	-135.87 ± 159.62*	-198.07 ± 167.64*
		Dual	-26.93 ± 18.86*	-41.73 ± 29.17*	-5.12 ± 15.40*
	Centre	Single	-164.39 ± 133.31*	-167.98 ± 150.09*	-154.92 ± 165.29*
		Dual	-19.98 ± 27.59*	-13.89 ± 23.99*	-23.22 ± 27.34*
	Bottom	Single	-77.55 ± 119.62*	-70.59 ± 130.98*	-53.05 ± 120.73
		Dual	-3.29 ± 38.36*	5.79 ± 27.57*	-3.14 ± 23.78
Spheres (Constant Size - 70mm)	Top	Single	-233.25 ± 174.06*	-96.18 ± 106.77*	-108.73 ± 158.88*
		Dual	-33.44 ± 15.35*	-40.36 ± 18.98*	-3.84 ± 24.62*
	Centre	Single	-155.76 ± 159.80*	-132.85 ± 123.58*	-60.61 ± 120.38*
		Dual	-7.00 ± 20.48*	-2.51 ± 20.87*	-8.36 ± 80.99*
	Bottom	Single	-81.21 ± 129.96*	-78.78 ± 112.20*	-41.74 ± 91.39
		Dual	-17.40 ± 20.84*	5.69 ± 48.60*	10.00 ± 18.27



(a) *gmp* placement in the z axis for cubes and spheres of all participants in 3 z planes (1400mm - 1600mm - 1800mm) using single view visual feedback



(b) *gmp* placement in the z axis for cubes and spheres of all participants in 3 z planes (1400mm - 1600mm - 1800mm) using dual view visual feedback

Figure 6.10: *gmp* placement in the z axis for cubes (black squares) and spheres (black circles) of all participants in 3 z planes (starting from top row in the order: 1400mm - 1600mm - 1800mm) using 6.10a: single view visual feedback and 6.10b: dual view visual feedback. Density heat maps indicate *gmp* placement across participants (red indicates higher density)

± 22.17) and spheres (-12.89mm ± 26.28)(see Table 6.7 [page 129]). This preference for the middle z plane (1600mm) can again be attributed to the convenience of this plane in terms of grasp planning (reaching and fine grasp adjustment, as it does not require as much flexion and extension of the forearm. For single view visual feedback, users showed a different preference as the lowest $GDisp_z$ means were found in the 1800mm z plane, and this is attributed to the position of this plane being at the extremity of the mean arm reach of users. The fact that users presented the highest accuracy in depth estimation in a plane that is not the furthest (1800mm) when provided with dual view visual feedback shows that accurate depth estimation is due to the feedback method used, whereas in single view visual feedback this accuracy is due to the test design and the biomechanical reach of users.

As shown in Figure 6.7j [page 119], lowest mean $GDisp_z$ under the dual view visual feedback condition was shown by users in the Centre positions for cubes (-0.95mm ± 25.25), and in the Top positions for spheres (-2.32mm ± 29.10) across all z planes. Again these interaction positions were convenient for users to perform freehand grasping within as they are easily reachable. However, accurate depth estimation in these positions did not necessarily lead to accurate size estimation of object sizes, especially for cubes as users preferred Bottom positions, not Centre, for accurate size estimation.

Alongside the x axis (see Figure 6.7e [page 119]), users showed the lowest $GDisp_z$ in Centre positions for cubes (-0.54mm ± 27.94), and in the Left positions for spheres (2.29mm ± 33.37) across all z planes.

Results from the two experiments in this study have shown that dual view visual feedback reduces the grasp variation problem that was presented in one visual feedback in Study 1 (Chapter 5), as users were more aware of their grasp shape and interaction using dual view visual feedback.

6.4.3.7 Results - Object Type

1400mm Z plane

Statistically significant differences between cubes and spheres in different positions were found in GAp ($\chi^2(1) = 3659$, $p < 0.01$), completion time ($\chi^2(1) = 340$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 179$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 257$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 51$, $p < 0.01$).

1600mm Z plane

Statistically significant differences between cubes and spheres in different positions were found in GAp ($\chi^2(1) = 3312$, $p < 0.01$), completion time ($\chi^2(1) = 282$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 17$, $p < 0.01$) and $GDisp_y$ ($\chi^2(1) = 661$, $p < 0.01$). No statistically significant difference was found in $GDisp_z$ ($\chi^2(1) = 4$, $p > 0.01$).

1800mm Z plane

Statistically significant differences between cubes and spheres in different positions were found in GAp ($\chi^2(1) = 911$, $p < 0.01$), completion time ($\chi^2(1) = 31$, $p < 0.01$), $GDisp_x$ ($\chi^2(1) = 78$, $p < 0.01$), $GDisp_y$ ($\chi^2(1) = 144$, $p < 0.01$) and $GDisp_z$ ($\chi^2(1) = 9$, $p < 0.01$)

6.4.3.8 Analysis - Object Type

In this section, findings for different object types (cubes and spheres) are reported per each z plane, and not for each individual position in this study to avoid repetition with results previously reported.

In GAp, users showed higher accuracy in grasping cubes than spheres in terms of matching GAp to object size across all positions, this was consistent across all z planes (see Table 6.8 [page 133]), and is also in alignment with findings for single view visual feedback where users also showed higher accuracy in grasping cubes than spheres. Using dual view visual feedback generally resulted in lower accuracy in GAp matching to objects size for both objects, with the exception of the 1800mm z plane where dual view visual feedback improved user performance in matching GAp to object size for both cubes and spheres across all positions.

In task completion time, users consistently showed lower mean completion times when grasping spheres than cubes (see Table 6.8 [page 133]). Users also showed less variation in completion times when grasping spheres, this was also found in Chapter 5 using single view visual feedback and can potentially be attributed to users preferring spheres over cubes on a perceptual level due to their shape and smaller size in this experiment, however this needs to be further analysed in future work in order to form a better understanding of the impact of perceiving object shapes on task completion times. As previously noted, completion times are not an ideal measure for accuracy in freehand grasping as users were more accurate in grasping cubes than spheres, and the longer completion times for cubes resulted in higher accuracy in terms of size matching potentially due to users spending longer times in fine grasp adjustments. Using dual view visual feedback resulted in higher completion times and variation for both objects (see Table 6.8 [page 133]), this was expected as users spent longer times adjusting their grasp using the additional feedback view.

In $GDisp_{\text{x}}$, users consistently showed lower mean $GDisp_{\text{x}}$ when grasping cubes than spheres (see Table 6.8 [page 133]). In addition, users also showed a lower variation in $GDisp_{\text{x}}$ for cubes than spheres, thus users were more confident in their gmp placement along the x axis when grasping cubes. Using dual view visual feedback resulted in lower $GDisp_{\text{x}}$ means for cubes, and higher $GDisp_{\text{x}}$ means for spheres (see Table 6.8 [page 133]).

In $GDisp_{\text{y}}$, users consistently showed lower mean $GDisp_{\text{y}}$ when grasping spheres than cubes, with he exception of the 1800mm z plane where a lower mean $GDisp_{\text{y}}$ was found for cubes than spheres (see Table 6.8 [page 133]). However, users showed lower $GDisp_{\text{y}}$ variation for cubes in all z planes, thus indicating that users were more confident in their gmp placement along the y axis, however this did not necessarily lead to higher accuracy (see Table 6.8 [page 133]). Using dual view visual feedback significantly lowered $GDisp_{\text{y}}$ means for both objects in comparison to single view visual feedback (see Table 6.8 [page 133]). In addition, dual view visual feedback also reduced the gmp deviation from the omp of objects (i.e. centre of mass of objects), thus

Table 6.8: Descriptive Statistics of Object Type (Mean ± SD). Significant differences ($p < 0.01$) between cubes and spheres in the dual view visual feedback condition are marked with (*)

1400mm Z Plane						
Object Type	View	GAp [mm]	Completion Time [s]	$GDisp_x$ [mm]	$GDisp_y$ [mm]	$GDisp_z$ [mm]
Cubes (80mm)	Single	81.86 ± 15.80	6.80 ± 4.38	27.20 ± 24.81	-4.73 ± 17.08	-13.30 ± 52.11
	Dual	71.91 ± 11.76 *	11.47 ± 7.14 *	23.00 ± 19.41 *	1.03 ± 12.59 *	20.55 ± 18.27*
Spheres (70mm)	Single	74.56 ± 11.97	5.86 ± 3.31	23.27 ± 29.28	-6.03 ± 21.59	-10.61 ± 56.19
	Dual	60.29 ± 15.60 *	9.56 ± 5.23 *	30.29 ± 33.52 *	-0.38 ± 20.09 *	28.02 ± 35.57*
1600mm Z Plane						
Object Type	View	GAp [mm]	Completion Time [s]	$GDisp_x$ [mm]	$GDisp_y$ [mm]	$GDisp_z$ [mm]
Cubes (80mm)	Single	79.94 ± 16.17	6.30 ± 5.29	30.39 ± 26.90	-9.80 ± 21.99	-58.75 ± 94.90
	Dual	74.82 ± 14.60*	11.43 ± 8.63*	26.22 ± 22.54*	4.21 ± 13.57*	3.34 ± 22.17
Spheres (70mm)	Single	73.87 ± 17.73	5.14 ± 2.63	24.32 ± 29.20	-10.33 ± 24.13	-51.60 ± 89.06
	Dual	60.91 ± 18.19*	10.41 ± 6.48*	27.55 ± 30.89*	0.29 ± 18.81*	5.07 ± 26.28
1800mm Z Plane						
Object Type	View	GAp [mm]	Completion Time [s]	$GDisp_x$ [mm]	$GDisp_y$ [mm]	$GDisp_z$ [mm]
Cubes (80mm)	Single	76.17 ± 22.90	6.80 ± 6.44	30.00 ± 33.75	-18.46 ± 36.22	-138.06 ± 159.08
	Dual	80.85 ± 23.30*	12.59 ± 9.87*	28.01 ± 35.84*	0.49 ± 20.45*	-14.61 ± 29.96*
Spheres (70mm)	Single	73.95 ± 23.50	5.21 ± 3.18	31.49 ± 42.48	-15.13 ± 39.36	-109.90 ± 144.09
	Dual	72.22 ± 25.40*	11.89 ± 8.14*	31.84 ± 39.00*	-1.90 ± 23.71*	-12.89 ± 39.27*

users were closer to the centre of the object when using dual view visual feedback (see Table 6.8 [page 133]).

In $GDisp_z$, users consistently showed lower mean $GDisp_z$ when grasping cubes than spheres, with the exception of the 1800mm z plane where a lower mean $GDisp_z$ was found for sphere than cubes (see Table 6.8 [page 133]. In addition, users showed lower variation and more confidence in depth estimation (gmp placement along the z axis) of cubes than spheres in all z planes. Using dual view visual feedback significantly lowered $GDisp_z$ means and variation for both objects in comparison to single view visual feedback, with the exception of the 1400mm where users showed better depth estimation under the single view visual feedback condition (see Table 6.8 [page 133]). Dual view visual feedback also significantly reduced the variation in depth estimation for both objects, thus users were more confident in depth estimation when they were provided with dual view visual feedback.

6.4.3.9 Usability Analysis

SUS average score for the different positions test was 64.50 (SD = 13.43). This rating of the dual view visual feedback is "OK and marginally acceptable" in experiment 2 (Bangor et al., 2009*a*). This rating was lower than that found for experiment 1 in this study, potentially due to the fact that users had to spend more time in locating objects in different positions in this experiment, unlike in experiment 1 where object position was constant. User comments using post test questionnaires below provide general subjective insights regarding their experience in grasping virtual objects using dual view visual feedback, however these insights may not be directly representative of user performance and accuracy during interaction as these subjective responses were not measured against performance in this work.

Out of 15 participants, 8 (53.3%) referred to look first to the frontal view while 6 (37.5%) focused their attention on the side view first, one user remained undecided.

To the question of which view was the most important for them, the opinion was divided into 9 (60.0%) users preferring to use the frontal view more, while 5 relied more on the side view (33.33%). With respect to which view was considered more important during the performance of the experiment, 11 (73.33%) users considered it to be the frontal view while 4 (26.66%) chose the side view. One user that choose the front view as most important commented saying "I just tried to establish the horizontal position of the object". Another user also commented "the front view is more important, but the side view seemed necessary for accuracy". Users that preferred the side view commented on their choice saying that the side view "gives more details about the depth information" and that using the side view "felt more natural". Users generally alternated between the two views in order to achieve an accurate grasp in this experiment.

On using the system again, 12 users (80.0%) will interact with the system again with dual view visual feedback. Regarding using two views one user commented that "the process might take longer, but the result seems more accurate". Another user further emphasised this point and commented that "it depends how accurate the grasp needs to be. For non-critical applications

the front view would be sufficient. For more critical applications, both views are required". These comments by users further show that interaction designers must take into account the accuracy speed trade-off that is associated with dual view visual feedback, as using this form is method can be largely dependent on the end goal of the application developed.
12 participants out of the 15 available had a specific approach for using dual visual feedback. Similar to experiment 1, the majority of users preferred using the front view first and then the side view. Some users also used a three step approach where they used front view, followed by the side and finally confirmed their grasp using the front view again.

Hypothesis - Revisited

$H_{2.2}$: using dual visual feedback in grasping virtual objects that change in position has no effect on a) grasp aperture and b) grasp displacement: **Rejected** as statistically significant results were found for the feedback method condition showing that dual visual feedback in grasping virtual objects that change in position has a significant effect on GAp, and on $GDisp$ in all axes (x, y and z).

6.5 Conclusions

This chapter presented a first study into the use of dual view visual feedback in an exocentric AR environment for assisting freehand grasping of virtual objects. The proposed measures of Grasp Aperture (GAp) and Grasp Displacement ($GDisp_{\text{xyz}}$) in Chapter 4 were used to quantify grasp ability and comparisons given against traditional single view visual feedback used in Chapter 5. This comprehensive study consisting of two experiments of the dual view visual feedback focused on mitigating the problems found in Chapter 5 that used the baseline single view visual feedback, namely grasp displacement in the x, y and axes ($GDisp_{\text{xy}}$), high displacement in the z axis ($GDisp_z$) and inaccurate object size estimation using GAp.
Findings in the two experiments illustrate that the proposed dual view visual feedback method in this study significantly improves Grasp Displacement in the x and y axes ($GDisp_{\text{xy}}$). Furthermore, user estimation of the object z position (the highest displacement found in the single view study) was significantly improved with the dual view feedback over single view feedback. This mitigation of displacement in the z axis was attributed to users increased awareness of their placement errors in the z axis via the additional side view feedback, thus allowing them to correct their grasp placement.
Similarities between the two feedback methods (single view and dual view) in user estimation of object size using GAp were also found. With single view feedback outperforming dual view visual feedback in matching GAp to object size. In Experiment 2 (i.e. changing position) participants were more focused on position change over object size, thus similar to single view feedback in Study 1 (Chapter 5), GAp varies less than expected using dual visual feedback, and was not proportional to object size. These findings are important when understanding how users

respond to different visual feedback views and noteworthy for future work developing freehand grasping systems. This study also shows that changing the visual feedback method does not improve size estimation using GAp, as it remains within a mean range of 60mm (SD = 31.28) to 70mm (SD = 31.09) across all participants and object types regardless of changes in object size and the feedback method employed.

Furthermore, completion time significantly increases using the proposed dual view visual feedback method in this study, thus even though the proposed feedback method in this study significantly improves spatial grasp placement, it results in longer completion times. This is attributed to participants repeatedly correcting their grasp posture for either aperture or position using the additional side view visual feedback. In addition, grasp variation that was present using single view visual feedback was reduced using dual view visual feedback. This indicates that enabling participants to visualise their hands using side and front views encourages more consistency in the grasp type.

Finally, from the usability analysis the following conclusions are drawn: the dual view visual feedback was rated as "good and acceptable" with a score 77 (SD = 16.45) for the object size experiment, while it was rated as "OK and marginally acceptable" for the object position experiment with a score of 64.5 (SD = 13.43). According to this, when the object position in the AR space changes for every test iteration participants found the use of dual view visual feedback more challenging due to the increased cognitive load of this proposed method. Finally, although there was a divided opinion in both experiments about which view is the most important, the majority of users concluded that they will interact again with the dual view visual feedback method, and consider this method more accurate and helpful for locating virtual objects in an AR environment.

In conclusion, this chapter measured the accuracy of freehand grasping in exocentric AR against virtual object size, position and type using a novel dual view visual feedback method. Findings in this study showed that the use of dual view visual feedback significantly improves depth perception and grasp spatial positioning in the x and y axes, thus mitigating some of the problems found in Chapter 5. However, this study has also shown that adding a second view camera for visual feedback does not improve the accuracy of GAp matching to object size in comparison to single view visual feedback. Moreover, significantly higher completion time was also present in both experiments in this study means that even though more accuracy can be achieved in spatial positioning of gmp in all axes, completion time increases a result. This shows that a speed-accuracy trade-off must be made before utilising dual view visual feedback. In Chapter 7 these two problems will be evaluated through implementing drop shadows as an additional feedback cue in the baseline environment that uses single view visual feedback (see Chapter 4) as an alternative to dual view visual feedback. Drop shadows will enable users to locate virtual objects faster by providing more information regarding the position of the virtual objects even before any interaction takes place. Whereas reverting back to single view visual feedback in the next study will improve virtual object size estimation using GAp given that results in this study

have shown that users show superior size estimation using single view visual feedback. The next chapter will revisit the problems presented in this study by quantifying the impact of this proposed feedback method on freehand grasping accuracy. In addition, the usability of drop shadows for freehand grasping will be discussed.

Chapter 7

Study 3: Drop Shadows for Freehand Grasping

7.1 Introduction

This chapter will present the third user study out of the four independent user studies in this work (see Figure 4.11 [page 64]) to assess the accuracy and usability of freehand grasping in exocentric AR. Findings in Chapter 6 have shown that freehand grasping accuracy can be improved using dual view visual feedback, through significantly improving grasp placement accuracy in the z axis that resulted in more accurate depth estimation of virtual objects. However dual view visual feedback has also presented key problems for freehand grasping, namely significantly high task completion times due to users spending longer times on correcting their grasp position using the additional feedback view, and higher inaccuracy in size estimation of virtual object sizes using GAp in comparison to single view visual feedback presented in Chapter 4. These problems could be attributed to the higher cognitive load associated with dual view visual feedback and can hinder the usability of freehand grasping in exocentric AR. This chapter will aim to address these problems by implementing drop shadows as an additional depth cue to allow users to locate virtual objects faster and more accurately.

7.1.1 Drop Shadows as a Depth Cue

Shadows are a crucial depth cue for humans in perceiving the 3D world around them and are useful in understanding size and position of virtual objects, the geometry of the surrounding environment and geometry of occluding objects or bodies (Hasenfratz et al., 2003). Current research offers strong evidence for the wide use, applications and impact of shadows as a depth cue in AR environments. For example, Diaz et al. (2017) recently evaluated the impact of different rendering effects such as shading, cast shadows, aerial perspective and texture on perceptual depth matching tasks, where users were instructed to match the position of various

virtual shapes to real world targets along the z axis using an HMD (HoloLens). Their work illustrated that shadows consistently had the largest impact on depth estimation in comparison to the other rendering features, whereas other cues such (e.g. shading, aerial perspective and texture) did not have a significant impact on depth estimation when used on their own, however still showed improvements when combined with other depth cues (such as shadows). This preference for shadows in their work highlighted the importance of virtual objects acting on the real world to achieve accurate depth estimation, this is evident by the fact that virtual (e.g. textures) and physical to virtual (e.g. shading) cues only improved depth estimation when used in conjunction with virtual to physical cues (drop and cast shadows). In addition, users also suggested that shadows were intuitive to use in interaction tasks as they directly understood their use in depth matching. Interestingly, this work also highlighted that users preferred drop shadows more than cast shadows even though cast shadows were more realistic in comparison to real world shadows, this was attributed to the effective human tolerance of imperfections in shadows as suggested in early studies in visual perception (Jacobson and Werner, 2004; Sattler et al., 2005).

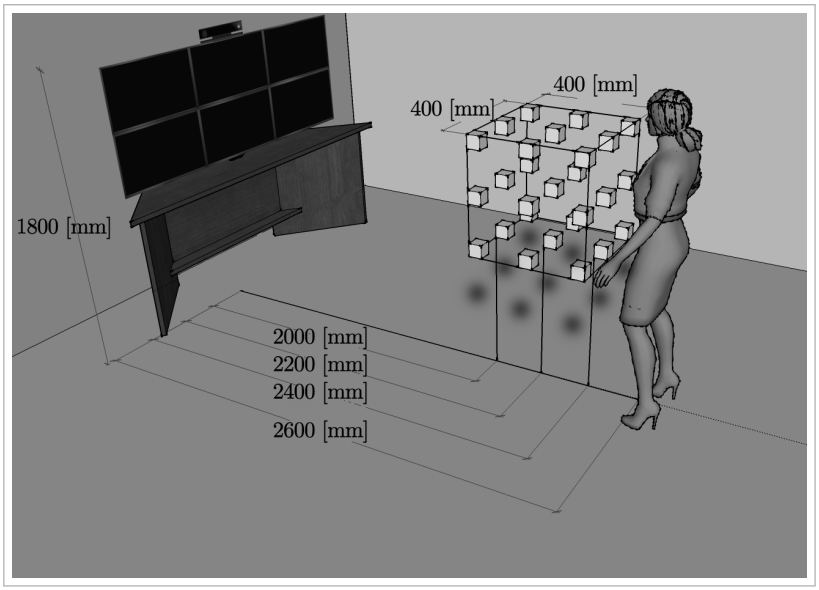
This chapter will aim to address the problems highlighted in Chapter 6 through presenting a first study into implementing drop shadows as an additional depth cue in exocentric AR for freehand grasping, this will potentially enable users to locate virtual objects faster, and can also offer another solution to inaccurate depth estimation in AR environments. The impact of drop shadows on freehand grasping accuracy will be measured using GAp and $GDisp$, and the usability of this proposed method will also be addressed using the standardised System Usability Scale (SUS). Section 7.2 [page 139] firstly outlines the design of the two experiments in this study in terms of the conditions under test, participants recruited, statistical model used and the experimental protocol. Section 7.3 [page 143] then discusses the data collected in the two experiments of this study that compare the two conditions: drop shadows and no drop shadows, and provides a comprehensive analysis of the interaction and usability results. Finally Section 7.4 [page 169] provides the conclusions drawn from this study and a summary of the key outputs that will be addressed in the following chapter.

7.2 Study Outline

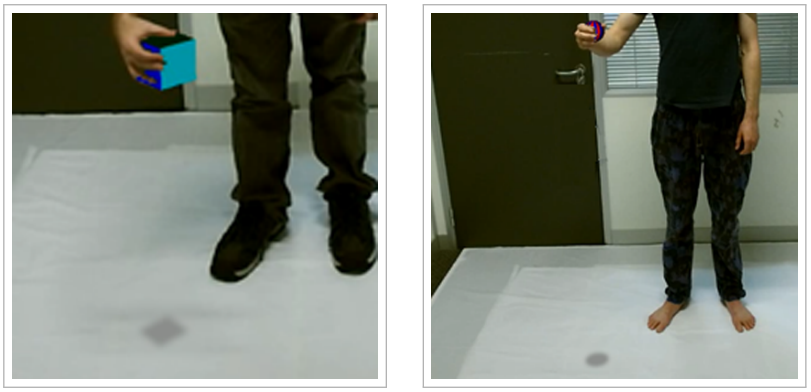
7.2.1 Design

Two experiments were conducted in this study using a slightly modified iteration of the baseline setup outlined in Chapter 4 (see Figure 7.1a [page 140]):

- Experiment 1 to quantify the influence of object position in x,y and z space and object type on grasp accuracy without using drop shadows
- Experiment 2 is a replication of experiment 1 to test the influence of object position in x,y and z space and object type on grasp accuracy using drop shadows



(a) Experiment Setup



(b) Interaction with virtual objects (cubes and spheres) with the implementation of drop shadows as a depth cue

Figure 7.1: Setup of the dual view visual feedback system developed

A 2 × 3 × 3 × 3 within-subjects design was used, with two primary conditions: drop shadows and no drop shadows. All new 15 participants took part in both conditions. Every permutation of position for both object types was randomly presented to participants to exclude potential learning effects. In total, each participant completed 27 (positions) × 2 (objects) = 54 trials and 810 grasps in total (54 trials × 15 participants). Findings from both experiments are compared to test the influence of drop shadows on grasp accuracy, thus in this chapter results from both experiments are presented together and not independently for each experiment to analyse the impact of the primary condition in this study that is drop shadows.

The system in this study is an adaptation of the baseline setup outlined in Chapter 4, where the only difference between the two systems was the positioning of the three z planes used (see Figure 7.1 [page 140]). In this study, participants stood 2500mm away from the sensor under controlled and constant lighting conditions, and the three z planes were placed at 2000mm, 2200mm and 2400mm distances away from the sensor instead of 1400mm, 1600mm and 1800mm

Table 7.1: Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor

Experiments 1 and 2				
Condition	**Levels**			
Object Position (x, y) [mm]		Left	Centre	Right
	Top	-400, 1650	0, 1650	400, 1650
	Centre	-400, 1250	0, 1250	400, 1250
	Bottom	-400, 850	0, 850	400, 850
	* 9 positions were repeated in each z plane (2000mm - 2200mm - 2400mm), resulting in 27 positions in total			
Drop Shadows	Off (Experiment 1) and On (Experiment 2)			
Object Type	Cube and Sphere			

as was implemented in studies 1 and 2 (see Figure 7.1a [page 140]). This adjustment was made to show the full ground in the environment in order for users to be able to see all the drop shadows on the floor in front of them as they usually do in real environments (see Figure 7.1 [page 140]). The floor was also covered with a white sheet to avoid a cluttered environment (floor in this case) and clearly distinguish drop shadows on the ground (see Figure 7.1b [page 140]). Drop shadows in this study were abstract, meaning that the shadows were not rendered based on the direction or intensity of a real world light source. Drop shadows in this study were essentially a second 3D object that was attached to the main object under test (i.e. cube or sphere). The objects were firstly modelled in Autodesk Maya, where the distance between the two objects was dependent on the physical distance between the main virtual object presented (i.e. cube or sphere) and the floor in a given task. Both objects were then rendered using OpenGL. Finally using OpenCV, the attached object (i.e. drop shadow) was blurred to soften the edges of the attached object providing the illusion of a shadow dropping from the main object in the test, and finally both objects were added to the final scene shown to users on the feedback monitor.

Conditions of both experiments are shown in Table 7.1[page 141], where experiment 2 is a replication of experiment 1 with the addition of drop shadows. The accuracy of a medium wrap grasp measured against the proposed metrics in this thesis; grasp aperture (GAp) and grasp displacement ($GDisp$) to test the impact of drop shadows on grasp accuracy. To represent the accuracy of a grasp independent of additional rendering, for both experiments, the baseline objects which have not undergone complex rendering and represent a simple abstract shape are used.

Hypotheses

$H_{3.1}$: adding drop shadows in freehand grasping of virtual objects that change in position has no effect on a) grasp aperture and b) grasp displacement
$H_{3.2}$: Adding drop shadows in freehand grasping of virtual objects that change in position has no effect on task completion time

7.2.2 Participants

30 participants ranged in age from 22 to 56 (M = 30.93, SD = 8.48), in arm length from 480mm to 660mm (M = 540.59, SD = 40.08), in hand size from 130mm to 210mm (M = 185.20, SD = 14.67), in height from 1558mm to 1940mm (M = 1729.83, SD = 84.45) and 8 were female and 22 male. Taking into account balance in hand size, arm length, gender, age and height, participants were separated into two groups of 15 for the two experiments.

7.2.3 Statistical Analysis

Due to the repeated measures design of this study and the format of the data collected being non-parametric and not normally distributed, statistical significance between the two independent groups in this study (drop shadows and no drop shadows) is tested using a non-parametric Wilcoxon signed-rank test (Wilcoxon and Wilcox, 1964) with an alpha of 1%.

7.2.4 Protocol

This study followed the baseline experiment protocol outlined in Section 4.2.4.2[page 62] prior to collection of data.
Participants underwent initial training of the medium wrap grasp on real and virtual objects and were given time to familiarise themselves with the modified environment in this study. The test coordinator explained the procedure between each block of tests (i.e cube and sphere), and participants were allowed to rest before the presentation of every object. Each experiment was formed of a 5 minutes training/instruction session, 10 minutes of grasping a cuboid object, 5 minutes break and 10 minutes of grasping a spherical object (order of virtual objects counterbalanced).
After completing the test, participants were asked to fill in two usability questionnaires and a set of questions regarding their interaction with the system. The usability of the system was evaluated by a user satisfaction test based on the System Usability Scale (SUS) (Brooke, 1996). This questionnaire consists of 10 items, which were evaluated by using a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Through feedback from this questionnaire, the ease of use and usability of drop shadows is evaluated. Finally, users were asked to answer a set of 8 close-ended questions. These questions were presented as a post-test questionnaire and participants commented on anything they considered related to the use of the system and drop shadows (see Appendix B [page 216]). Questions/Statements were:

1. I found it easy to locate and successfully grasp objects
2. I have noticed that the virtual objects changed position in the x, y and z axis
3. Did you suffer from fatigue or pain?
4. Which of the two objects did you find easier to interact with?
5. I have noticed the drop shadows changed in position in the x, y and z axes depending on the object's position [drop shadows condition]
6. I used the drop shadows to locate the virtual objects presented [drop shadows condition]
7. I found the drop shadows useful in accurately locating virtual objects [drop shadows condition]
8. Which depth cue did you find to be more useful in locating virtual objects? Occlusion - Drop shadows - Both? [drop shadows condition]

7.2.5 Procedure

For the two experiments (with and without drop shadows), participants were instructed to accurately locate and match their grasp aperture and position to the size and position of the virtual object in the shortest time possible. 27 different positions in all axes (x, y and z) are used (see Table 7.1 [page 141]), covering a working range of 400mm from participants (see Figure 7.1a [page 140]). The object sizes that had the lowest mean difference between GAp and object size found in previous studies in this thesis were chosen (80mm for cubes and 70mm for spheres) and were unchanged throughout the two experiments in this study.
Before interaction, an object (cube or sphere) appeared to participants on the feedback monitor, each object had 27 different positions. A countdown of 5 seconds followed by an auditory cue was used as an indicator for participants to start grasping the object
During the experiment, all participants were instructed to verbally inform the test coordinator that they are satisfied with the grasp they have performed, and maintain the grasp for 5 seconds while the measurements are stored.

7.3 Results

7.3.1 Results - Grasp Aperture (GAp)

2000m Z plane

A statistically significant difference in GAp between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.41 $\times 10^7$, p < 0.01) and spheres (Z = 1.71 $\times 10^7$, p < 0.01)

2200m Z plane

A statistically significant difference in GAp between the drop shadows and no drop shadows conditions in different positions was found for grasping cubes (Z = 1.53 $\times 10^7$, p < 0.01) and spheres (Z = 1.60 $\times 10^7$, p < 0.01)

2400m Z plane

A statistically significant difference in GAp between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.40 $\times 10^7$, p < 0.01) but not spheres (Z = 1.43 $\times 10^7$, p > 0.01)

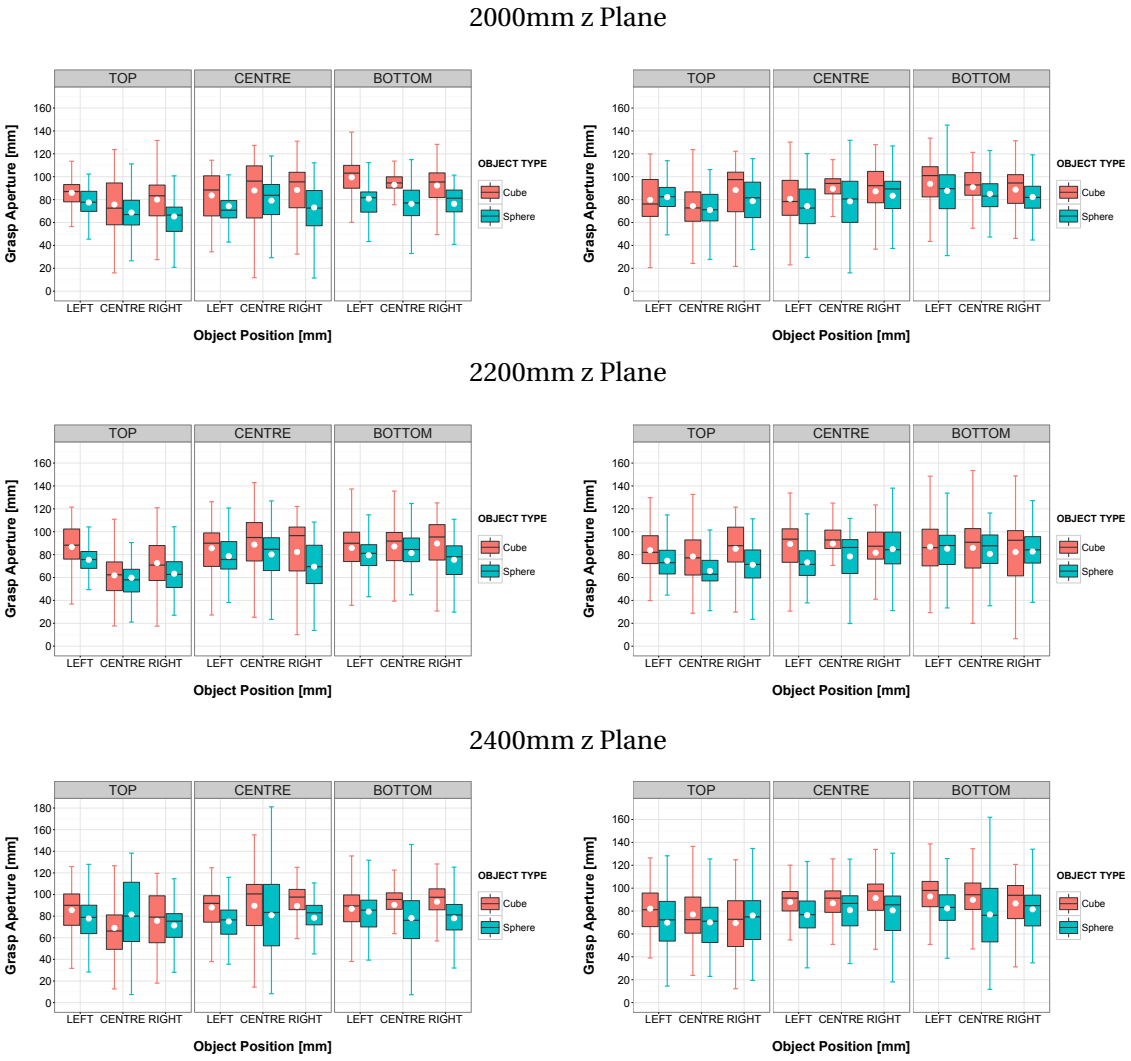
7.3.2 Analysis - Grasp Aperture (GAp)

Users have overestimated object sizes in both the drop shadows and no drop shadows conditions for both objects, this overestimation of object size was consistent across all z planes, where mean underestimation of object size only occurred in 17 out of the 108 trials in all conditions under test (see Table 7.2 [page 145]). Despite the constant size of virtual objects in this study, high variation in GAp was shown by users in grasping both objects across all z planes. For cubes, GAp ranged from 69.49mm ± 17.75 to 93.77mm ± 20.67 in the no drop shadows condition, whereas a wider range from 61.89mm ± 17.75 to 99.45mm ± 20.67 was shown in the drop shadows condition. In contrast for spheres, GAp ranged from 65.74mm ± 14.15 to 87.69mm ± 23.38 in the drop shadows condition, whereas a narrower range from 59.78mm ± 21.45 to 84.17mm ± 21.04 was shown in the drop shadows condition. This high variation in GAp is attributed to the lack of tactile feedback in the hand, that is key in the process of finely adjusting grasp parameters such as the GAp and is comparable to Studies 1 and 2 in this work (see Chapters 5 and 6). Performing a certain GAp is usually dictated partly by the object constraints in a given task, however due to this lack of tactile feedback in freehand grasping, object constraints such as the size do not necessarily present a GAp that is proportional to the size of the virtual object, even if the size of the virtual object is altered significantly as illustrated in the first and second studies in this thesis.

Users presented the highest accuracy in terms of matching their GAp to object sizes in the 2200mm z plane for cubes in both conditions (drop shadows and no drop shadows), with a mean overestimation of 2.26mm ± 24.43 in the drop shadows condition, and a mean overestimation of 4.77mm ± 27.75 in the no drop shadows condition across all positions (see Figures 7.2a [page 146] and 7.2b [page 146]). Similarly for spheres, users also presented the highest accuracy in GAp in the 2200mm z plane in the drop shadows condition with a mean overestimation of 3.67mm ± 19.95 across all positions, and in the 2400mm z plane in the no drop shadows condition with a mean overestimation of 7.19mm ± 22.82. However, this was comparable to the mean overestimation in the 2200mm z plane for spheres (7.37mm ± -48.33) as the difference between the two planes was not significant (p > 0.01). This preference for the middle z plane by users is in alignment with findings reported in the two previous studies in this thesis (Chapters 5 and 6), and is potentially due to the convenience of its spatial position, as this particular plane does not require as much flexion and extension of the forearm as the other two planes that

Table 7.2: Descriptive Statistics of GAp (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*)

2000mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	No Drop Shadows	79.84 ± 20.20	74.43 ± 19.78	88.36 ± 23.45
		Drop Shadows	85.96 ± 11.21*	75.74 ± 23.69	80.03 ± 18.73*
	Centre	No Drop Shadows	80.72 ± 21.15	89.49 ± 14.57	87.20 ± 24.98
		Drop Shadows	83.67 ± 19.34	87.95 ± 24.31*	88.41 ± 26.20
	Bottom	No Drop Shadows	93.77 ± 22.52	90.83 ± 17.17	88.71 ± 15.90
		Drop Shadows	99.45 ± 20.67*	92.95 ± 15.96	92.31 ± 17.37*
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	82.20 ± 14.93	70.91 ± 17.46	78.66 ± 18.33
		Drop Shadows	77.47 ± 11.98*	68.70 ± 19.22*	65.25 ± 20.09*
	Centre	No Drop Shadows	74.35 ± 17.23	78.46 ± 21.66	83.40 ± 19.84
		Drop Shadows	74.26 ± 15.25	79.16 ± 20.49	73.15 ± 19.32*
	Bottom	No Drop Shadows	87.69 ± 23.38	85.05 ± 20.31	82.31 ± 20.65
		Drop Shadows	80.83 ± 19.75*	76.35 ± 19.36*	76.38 ± 16.53*
2200mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	83.87 ± 16.80	78.40 ± 18.69	85.12 ± 22.98
		Drop Shadows	86.66 ± 19.58*	61.89 ± 17.75*	72.64 ± 21.39*
	Centre	No Drop Shadows	89.22 ± 20.94	89.55 ± 20.86	81.69 ± 25.58
		Drop Shadows	85.53 ± 20.65*	88.75 ± 25.01	82.35 ± 30.80*
	Bottom	No Drop Shadows	86.86 ± 23.03	85.91 ± 22.44	82.28 ± 25.23
		Drop Shadows	85.73 ± 23.11	87.11 ± 22.29	89.67 ± 22.80*
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	74.65 ± 20.13	65.74 ± 14.15	71.11 ± 17.46
		Drop Shadows	75.44 ± 15.59	59.78 ± 21.45*	63.34 ± 14.43*
	Centre	No Drop Shadows	73.17 ± 16.92	78.38 ± 22.24	84.76 ± 23.18
		Drop Shadows	78.60 ± 15.88*	80.12 ± 19.74	69.48 ± 21.18*
	Bottom	No Drop Shadows	85.18 ± 27.24	80.63 ± 21.63	82.71 ± 20.91
		Drop Shadows	79.43 ± 18.63*	81.37 ± 21.58	75.50 ± 17.07*
2400mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	82.11 ± 17.73	77.01 ± 28.28	69.49 ± 23.42
		Drop Shadows	85.55 ± 19.08*	69.17 ± 29.49*	75.66 ± 25.66*
	Centre	No Drop Shadows	87.80 ± 16.08	86.81 ± 19.09	91.36 ± 20.10
		Drop Shadows	88.01 ± 21.26	89.63 ± 27.97*	89.40 ± 25.38
	Bottom	No Drop Shadows	92.74 ± 21.32	89.79 ± 21.92	86.65 ± 21.39
		Drop Shadows	86.78 ± 20.02*	90.37 ± 18.43	93.35 ± 17.65*
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	69.84 ± 22.21	70.11 ± 24.95	76.14 ± 24.43
		Drop Shadows	77.78 ± 18.30*	81.52 ± 30.92*	71.36 ± 18.47
	Centre	No Drop Shadows	76.44 ± 20.84	80.76 ± 17.21	80.58 ± 27.16
		Drop Shadows	75.09 ± 20.52	80.85 ± 37.71	78.48 ± 18.41
	Bottom	No Drop Shadows	82.33 ± 20.66	77.02 ± 29.25	81.53 ± 21.85
		Drop Shadows	84.17 ± 21.04	78.21 ± 29.27	78.00 ± 16.51*



(a) Drop shadows

(b) No drop shadows

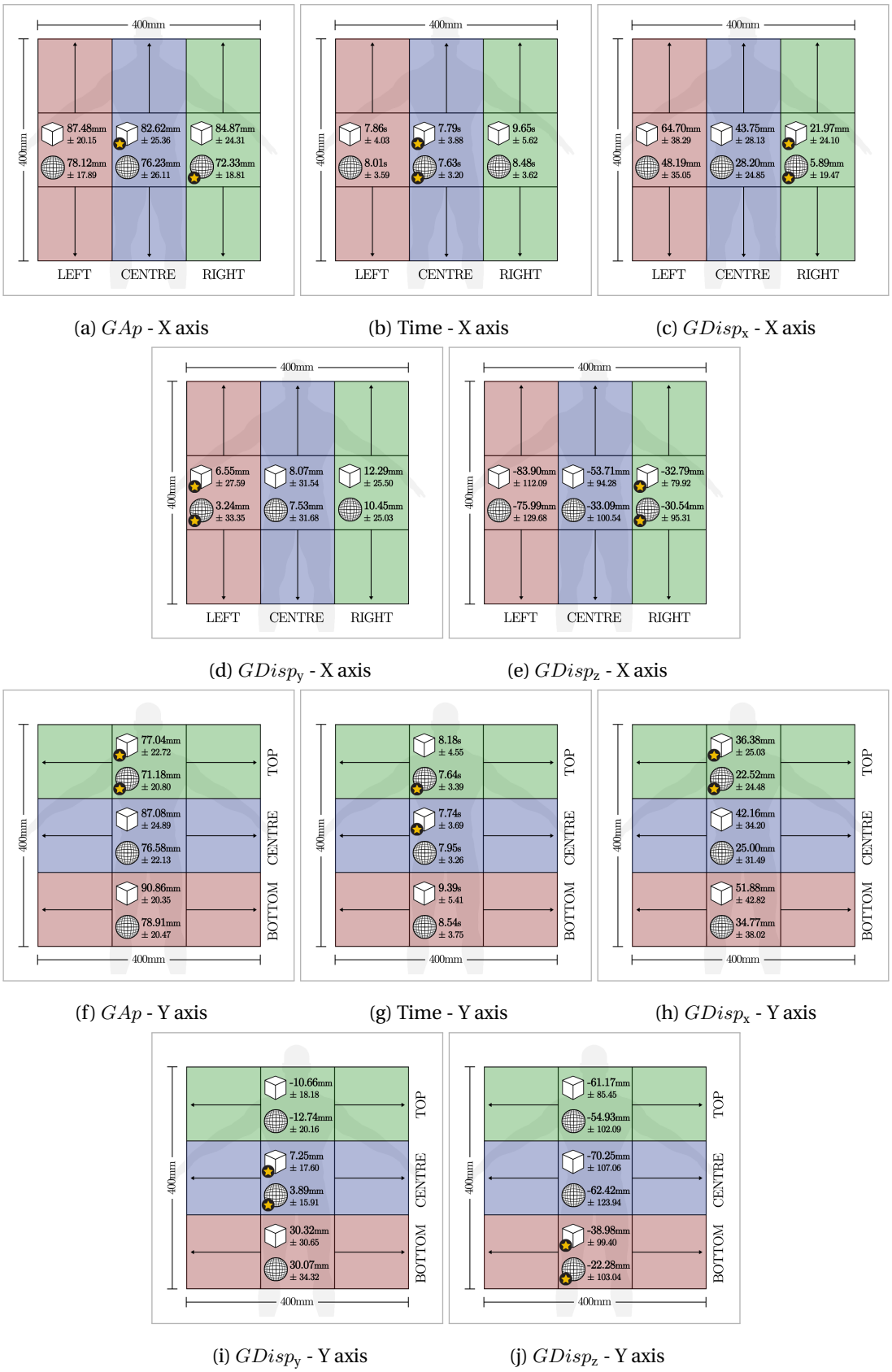
Figure 7.2: GAp for different object positions in the three z planes in this experiment (2000mm, 2200mm and 2400mm). 7.2a: Drop shadows used. 7.2b: No drop shadows used. White points on boxplots indicate the mean GAp across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

represent the extremities of the working range within the environment, and hence it is more comfortable to grasp within than the furthest (2000mm) and closest (2400mm) z planes that can be more physically demanding for users in terms of arm reach and fine grasp adjustments.

As shown in Figures 7.3f [page 148] and 7.4f [page 149], users showed the highest accuracy in matching GAp to object size in the Top positions alongside the y axis for cubes with a mean underestimation of -2.97mm $\pm$ 22.72 in the drop shadows condition, and a mean underestimation of -0.15mm $\pm$ 22.19 in the no drop shadows condition. Similarly for spheres, users also showed the highest accuracy in matching GAp to object size in the Top positions with a mean overestimation of 1.18mm $\pm$ 20.80 in the drop shadows condition, and a mean overestimation of 3.26mm $\pm$ 20.24 in the no drop shadows condition.

Alongside the x axis (see Figures 7.3a [page 148] and 7.4a [page 149]), users showed the highest accuracy in matching GAp to object size in the Centre positions for cubes with a mean overestimation of 2.62mm $\pm$ 25.36 in the drop shadows condition, and in the Right positions in the no drop shadows with a mean overestimation of 4.54mm $\pm$ 23.52. For spheres, Centre positions showed the highest accuracy in GAp with a mean overestimation of 2.33mm $\pm$ 18.81 in the drop shadows condition, and in the right positions with a mean overestimation of 6.34 $\pm$ 22.16 in the no drop shadows condition. Similar to findings in Study 1 (Chapter 5), this shows that combinations of these interaction regions (Top/Centre and Right/Centre) yield the highest accuracy in GAp matching to object sizes, and highlights the importance of placing virtual objects in positions that are easily accessible using the dominant hand, regardless of the depth cues used.

In terms of accuracy, users performed better in matching GAp to object size in the no drop shadows condition than the drop shadows condition in the majority of individual positions across the three z planes for both objects (see Table 7.2 [page 145]), with the exception of 23 out of the 54 trials in this study where the drop shadows outperformed the no drop shadows condition, however only 13 trials showed statistically significant differences (p < 0.01). This shows that adding drop shadows as an additional depth cue does not necessarily improve GAp estimation as users performed better without drop shadows in the majority of positions. This can potentially be attributed to additional cognitive load in the process of grasping, as users in the drop shadows condition users were also focused on the spatial positioning of their grasp in relation to the object and its shadow alongside GAp estimation (see Figures 7.2a [page 146] and 7.2b [page 146]). Wherein the no drop shadows condition users were solely focused on GAp estimation without the additional need of accurately placing their grasp in relation to the object. A similar finding was also present in Study 2 (Chapter 6) where an additional depth cue (dual view) was used, where users did not show improvements in GAp using dual view visual feedback, possibly due to being more focused on correcting the position of their grasp using the second view. Understanding the impact of cognitive load on freehand grasping of virtual objects in AR can be a valuable route for future work to further understand the limitations of additional depth cues to assist users in this form of natural hand interaction.



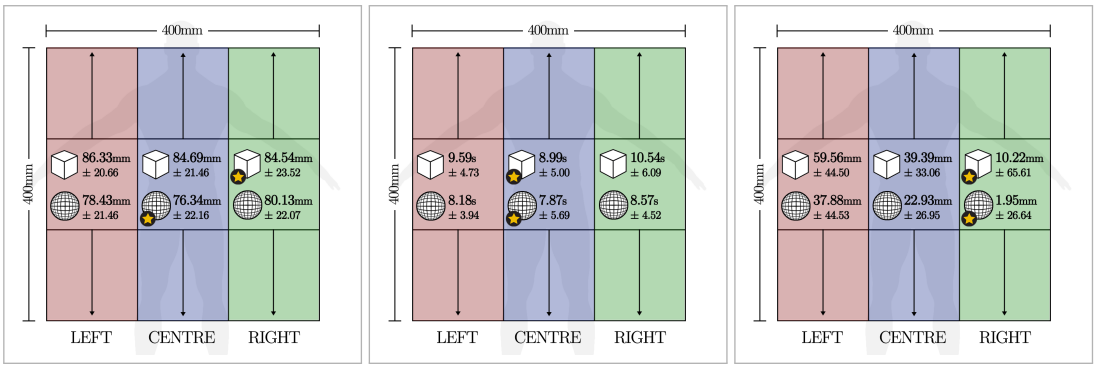
(a) GAp - X axis (b) Time - X axis (c) $GDisp_x$ - X axis

(d) $GDisp_y$ - X axis (e) $GDisp_z$ - X axis

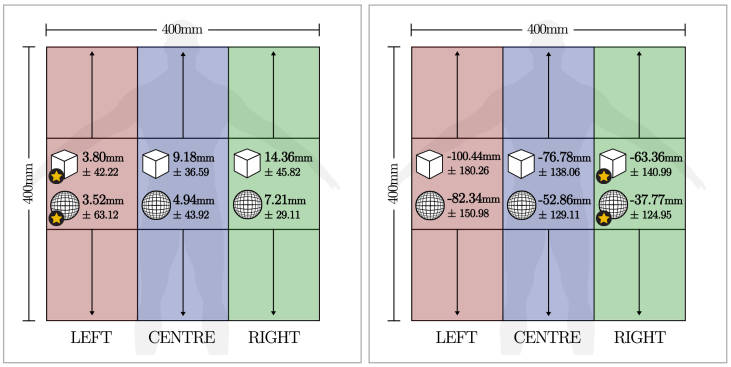
(f) GAp - Y axis (g) Time - Y axis (h) $GDisp_x$ - Y axis

(i) $GDisp_y$ - Y axis (j) $GDisp_z$ - Y axis

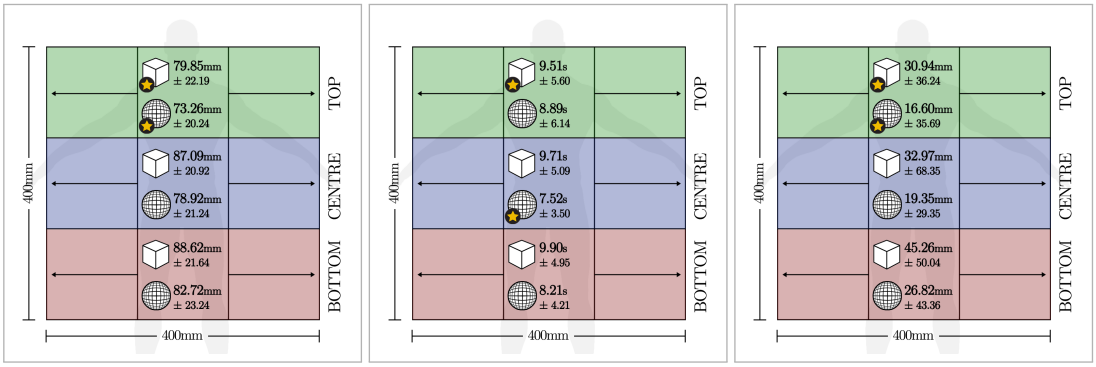
Figure 7.3: **Drop Shadows Condition:** Optimal interaction regions for users across all z planes in all the measurement used in this work to assess grasp accuracy. X axis refers to Left - Centre - Right positions. Y axis refers to Top - Centre - Bottom positions. Values presented are Means ± SD of each measurement. Most accurate/quickest region is marked with a star



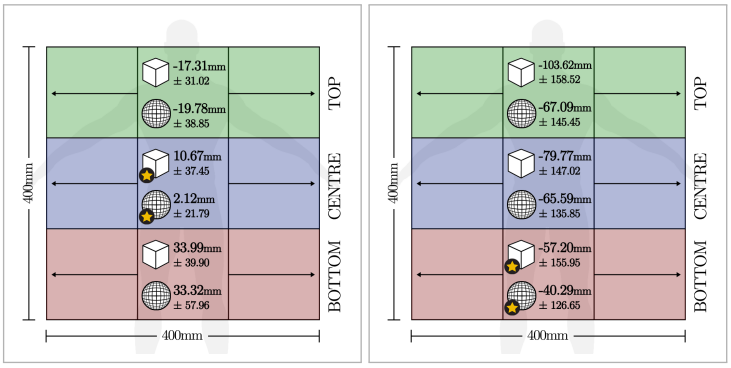
(a) GAp - X axis (b) Time - X axis (c) $GDisp_x$ - X axis



(d) $GDisp_y$ - X axis (e) $GDisp_z$ - X axis



(f) GAp - Y axis (g) Time - Y axis (h) $GDisp_x$ - Y axis



(i) $GDisp_y$ - Y axis (j) $GDisp_z$ - Y axis

Figure 7.4: **No Drop Shadows Condition:** Optimal interaction regions for users across all z planes in all the measurement used in this work to assess grasp accuracy. X axis refers to Left - Centre - Right positions. Y axis refers to Top - Centre - Bottom positions. Values presented are Means ± SD of each measurement. Most accurate/quickest region is marked with a star

7.3.3 Results - Completion Time

2000mm Z plane

A statistically significant difference in completion time between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.72 $\times 10^7$, p < 0.01) and spheres (Z = 1.55 $\times 10^7$, p < 0.01)

2200mm Z plane

A statistically significant difference in completion time between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.63 $\times 10^7$, p < 0.01) and spheres (Z = 1.37 $\times 10^7$, p < 0.01)

2400mm Z plane

A statistically significant difference in completion time between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.68 $\times 10^7$, p < 0.01) and spheres (Z = 1.29 $\times 10^7$, p < 0.01)

7.3.4 Analysis - Completion Time

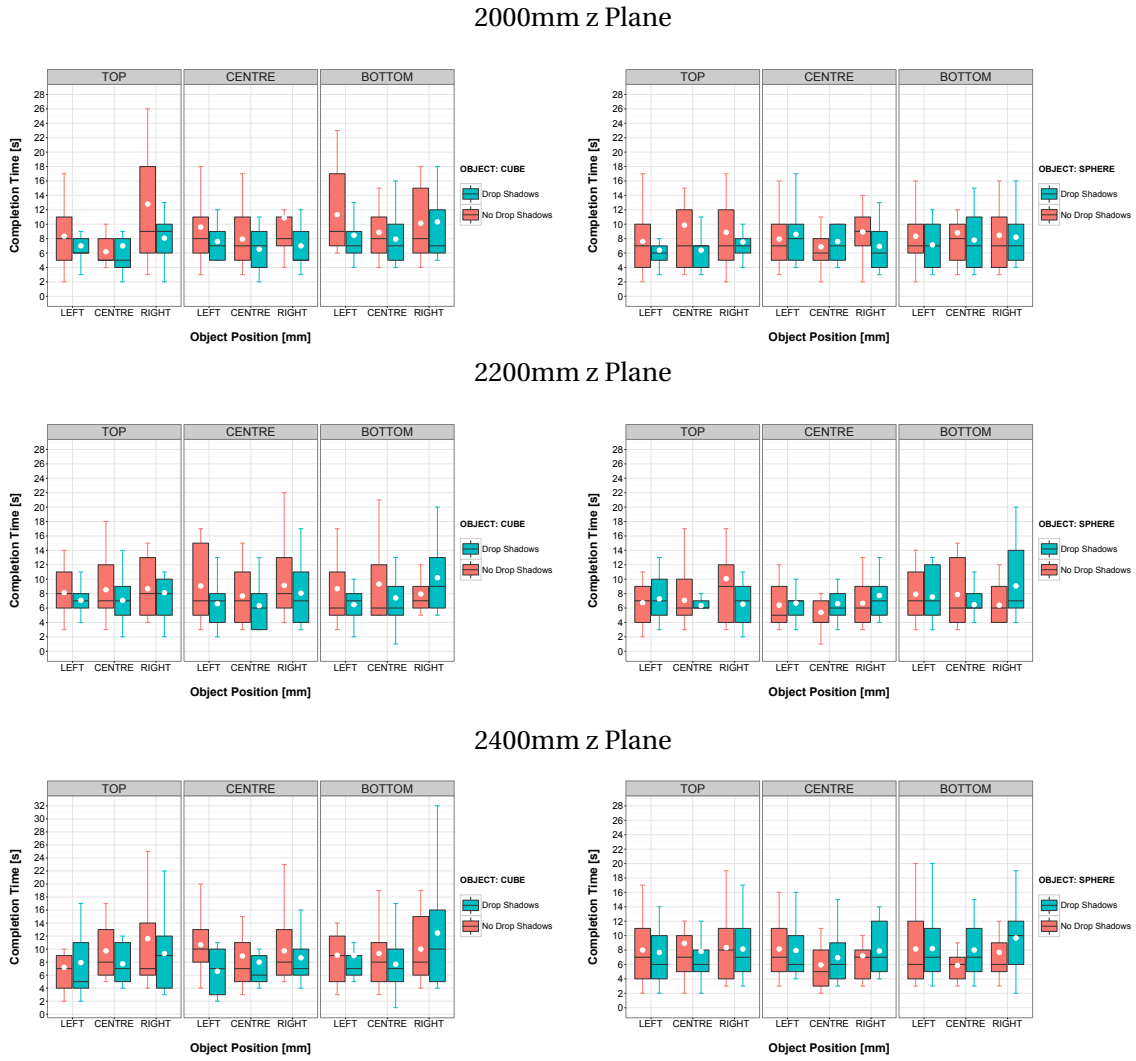
Mean completion time ranged from 6.63s ± 1.85 to 13.28s ± 9.41 for cubes in the no drop shadows condition, and a lower completion time range was found in the drop shadows condition from 6.86s ± 3.08 to 13.05s ± 9.16. Similarly for spheres, mean completion time ranged from 5.87s ± 2.85 to 10.57s ± 7.22 in the no drop shadows condition, and from 6.78s ± 1.22 to 10.26s ± 4.01 in the drop shadows condition. This high variation in completion time across different object types and z planes is potentially reflective of the individual differences between users, as choosing a grasp strategy in perceptual tasks can be impacted by individual differences in perception and experience thus resulting in varying completion times between users.

Users again consistently presented the shortest completion time in the 2200mm z plane for cubes and spheres in both conditions (drop shadows and no drop shadows) across all positions, where the drop shadows condition presented shorter completion times in this plane for both objects (see Figures 7.5a [page 152] and 7.5b [page 152]). For cubes, mean completion time in this particular z plane was 7.95s ± 3.82 in the drop shadows condition and 9.01s ± 4.31 in the no drop shadows condition. For spheres, mean completion time was 7.62s ± 2.96 in the drop shadows condition, and 7.70s ± 3.96 in the no drop shadows condition. This preference for the 2200mm z plane can again be attributed to the convenience of its position in relation to the biomechanical reach of users, as this plane is not as physically demanding as the furthest and closest z planes.

As shown in Figures 7.3g [page 148] and 7.4g [page 149], users also showed shortest mean completion times in the Centre positions alongside the y axis across all z planes for cubes with a mean completion time of 7.74s ± 3.69 in the drop shadows condition, and in Top positions in the no drop shadows condition with a longer mean completion time of 9.51s ± 5.60. For spheres, users showed the shortest mean completion time across all z planes in the Top positions with a

Table 7.3: Descriptive Statistics of Task Completion Time (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*)

2000mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	No Drop Shadows	8.81 ± 4.31	6.63 ± 1.85	13.28 ± 9.41
		Drop Shadows	7.42 ± 2.85	7.54 ± 5.33	8.65 ± 3.08
	Centre	No Drop Shadows	10.03 ± 4.88	8.32 ± 3.65	11.17 ± 7.35
		Drop Shadows	8.06 ± 2.42	7.01 ± 2.67	7.48 ± 3.40
	Bottom	No Drop Shadows	11.88 ± 5.26	9.40 ± 3.05	10.68 ± 5.01
		Drop Shadows	8.92 ± 3.70	8.44 ± 3.36	10.76 ± 6.26
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	8.06 ± 4.33	10.31 ± 10.14	9.30 ± 4.43
		Drop Shadows	6.97 ± 2.74	6.96 ± 2.51	8.21 ± 2.26
	Centre	No Drop Shadows	8.39 ± 3.70	7.32 ± 2.87	9.41 ± 3.70
		Drop Shadows	9.10 ± 3.61*	8.22 ± 3.56	7.40 ± 2.84*
	Bottom	No Drop Shadows	8.81 ± 4.37*	9.36 ± 5.03	9.12 ± 4.87
		Drop Shadows	7.69 ± 3.06*	8.36 ± 4.00	8.70 ± 3.54
2200mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	8.54 ± 3.19	8.96 ± 3.81	9.21 ± 3.60
		Drop Shadows	7.56 ± 1.86	7.37 ± 3.43	8.70 ± 5.24
	Centre	No Drop Shadows	9.52 ± 4.73	8.21 ± 3.90	9.44 ± 4.68
		Drop Shadows	6.99 ± 2.79	6.86 ± 3.08	8.55 ± 4.51
	Bottom	No Drop Shadows	9.16 ± 5.11	9.82 ± 5.64	8.27 ± 3.14
		Drop Shadows	6.92 ± 1.96	7.90 ± 3.89	10.75 ± 4.46
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	7.28 ± 2.47	7.66 ± 3.43	10.57 ± 7.42
		Drop Shadows	7.73 ± 3.03	6.78 ± 1.22	6.94 ± 2.62
	Centre	No Drop Shadows	7.00 ± 2.39	5.87 ± 2.85	7.18 ± 2.82
		Drop Shadows	7.19 ± 2.03	7.15 ± 1.95	8.24 ± 3.49
	Bottom	No Drop Shadows	8.45 ± 3.15	8.35 ± 4.16	6.91 ± 2.33
		Drop Shadows	7.95 ± 3.30	7.07 ± 2.03	9.56 ± 4.48
2400mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	7.74 ± 3.85	10.28 ± 5.07	12.10 ± 7.67
		Drop Shadows	8.37 ± 7.32	8.30 ± 2.79	9.73 ± 5.63
	Centre	No Drop Shadows	11.08 ± 4.37	9.48 ± 5.83	10.14 ± 4.54
		Drop Shadows	7.01 ± 3.19	8.52 ± 5.10	9.15 ± 4.31
	Bottom	No Drop Shadows	9.56 ± 5.06	9.80 ± 5.77	10.53 ± 4.82
		Drop Shadows	9.53 ± 5.74	8.22 ± 3.96	13.05 ± 9.16
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	8.55 ± 4.06	9.43 ± 9.04	8.83 ± 4.17
		Drop Shadows	8.26 ± 5.02	8.26 ± 4.53	8.66 ± 4.15
	Centre	No Drop Shadows	8.51 ± 4.26	6.41 ± 3.63	7.57 ± 3.43
		Drop Shadows	8.37 ± 3.72	7.52 ± 3.52	8.37 ± 3.44
	Bottom	No Drop Shadows	8.53 ± 5.36	6.15 ± 1.87	8.22 ± 4.24
		Drop Shadows	8.85 ± 4.34	8.38 ± 3.41	10.26 ± 4.01



(a) Cubes

(b) Spheres

Figure 7.5: Task completion times for different object positions in the three z planes in this study (2000mm, 2200mm and 2400mm). 7.5a: Cubes. 7.5b: Spheres. White points on boxplots indicate the mean completion time across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

mean completion time of 7.64s ± 3.39 in the drop shadows condition, and in the Centre positions in the no drop shadows condition with a shorter mean completion time of 7.52s ± 3.50 in the no drop shadows condition.

Alongside the x axis (see Figure 7.3b [page 148] and 7.4b [page 149]), users showed the shortest completion time in the Centre positions for cubes with mean completion times of 7.79s ± 3.88 in the drop shadows condition and 8.99s ± 4.60 in the no drop shadows condition. Centre positions also showed the shortest completion times for spheres across all conditions (drop shadows and no drop shadows), with a mean completion time of 7.63s ± 3.20 in the drop shadows condition and 7.87 ± 5.69 in the no drop shadows condition. These interaction regions across the x (Centre and Right) and y (Centre and Top) axis are usually preferred by users as they are easily accessible using grasp to reach movements, and also feel more natural to interact within using the right dominant hand.

Use of drop shadows significantly reduced completion time in the majority of positions under test (see Table 7.3 [page 151]), where the drop shadows condition outperformed the no drop shadows condition in terms of completion time in 37 out of the 54 trials in this study with 29 out of the 37 showing statistical significance (p < 0.01). This shows that using drop shadows as an additional depth cue can positively impact task completion time in freehand grasping, as drop shadows ease the process of locating virtual objects in different spatial positions and especially in different z planes. This impact on completion time is important in freehand grasping, as drop shadows can guide users to the position of the virtual object without spending the majority of interaction time in reaching and searching for objects which can negatively impact usability and potentially cause fatigue. In addition, a negative experience in the reaching phase of a grasp (i.e. spending a long time in locating virtual objects) can potentially hinder the performance in the next phases of a grasp.

7.3.5 Results - Grasp Displacement ($GDisp$)

2000mm Z plane

A statistically significant difference in $GDisp_{\text{x}}$ between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.26 $\times 10^7$, p < 0.01) and spheres (Z = 1.33 $\times 10^7$, p < 0.01).

A statistically significant difference in $GDisp_{\text{y}}$ between the drop shadows and no drop shadows conditions in different positions was found for spheres (Z = 1.30 $\times 10^7$, p < 0.01), but not cubes (Z = 1.42 $\times 10^7$, p > 0.01).

A statistically significant difference in $GDisp_{\text{z}}$ between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.35 $\times 10^7$, p > 0.01) and spheres (Z = 1.30 $\times 10^7$, p < 0.01).

2200mm Z plane

A statistically significant difference in $GDisp_{\text{x}}$ between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.29 $\times 10^7$, p < 0.01) and spheres (Z =

1.35 $\times 10^7$, p < 0.01)

A statistically significant difference in $GDisp_y$ between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.39 $\times 10^7$, p > 0.01) and spheres (Z = 1.38 $\times 10^7$, p < 0.01).

A statistically significant difference in $GDisp_z$ between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.38 $\times 10^7$, p > 0.01) and spheres (Z = 1.38 $\times 10^7$, p < 0.01).

2400mm Z plane

A statistically significant difference in $GDisp_x$ between the drop shadows and no drop shadows conditions in different positions was found for cubes (Z = 1.38 $\times 10^7$, p < 0.01) and spheres (Z = 1.25 $\times 10^7$, p < 0.01)

A statistically significant difference in $GDisp_y$ between the drop shadows and no drop shadows conditions in different positions was found for spheres (Z = 1.39 $\times 10^7$, p < 0.01), but not cubes (Z = 1.49 $\times 10^7$, p > 0.01).

A statistically significant difference in $GDisp_z$ between the drop shadows and no drop shadows conditions in different positions was found for spheres (Z = 1.40 $\times 10^7$, p < 0.01), but not cubes (Z = 1.47 $\times 10^7$, p > 0.01).

7.3.6 Analysis - Grasp Displacement ($GDisp$)

***GDisp*$_x$**

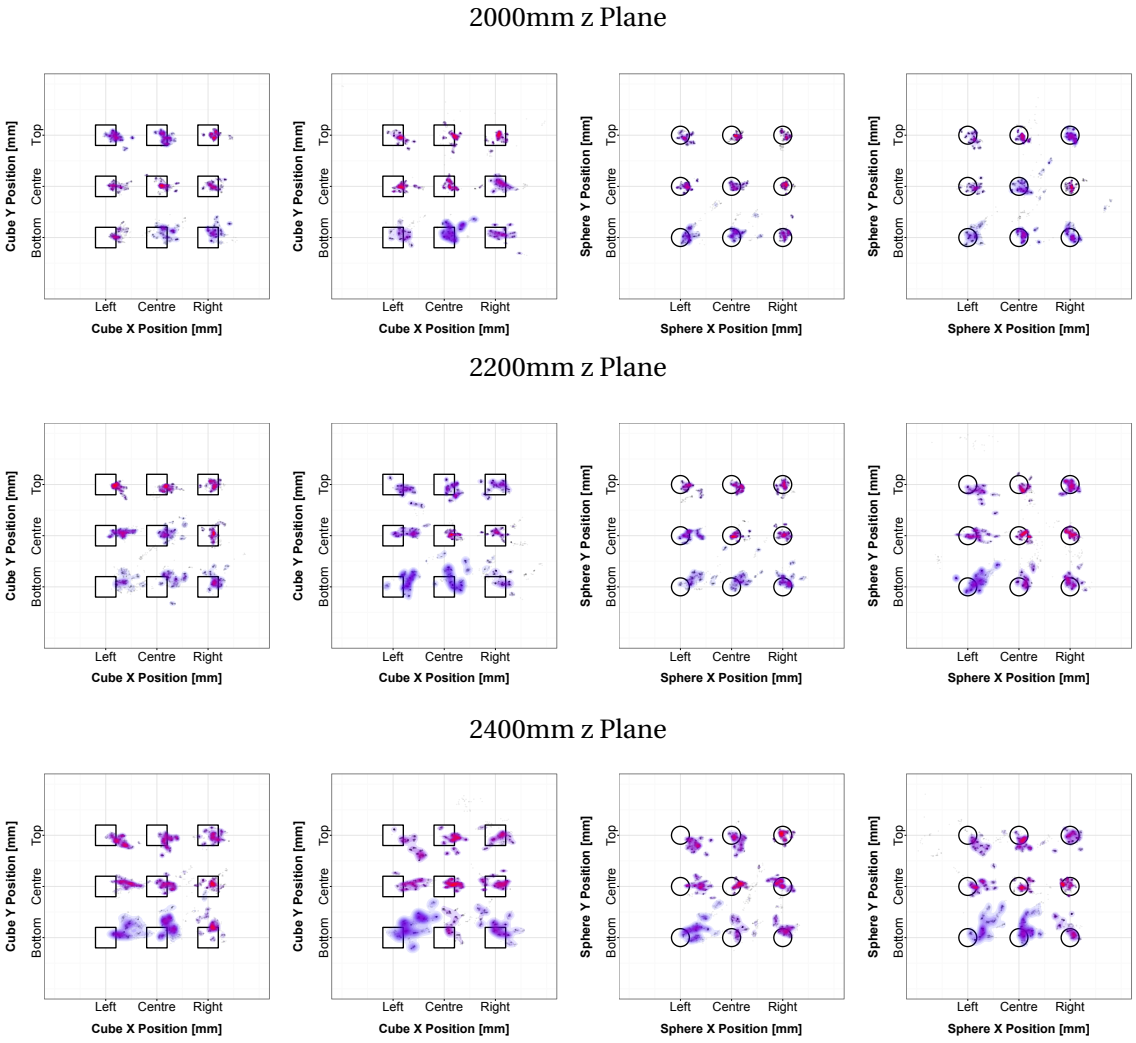
$GDisp_x$ showed the second highest displacement in this study in both conditions (drop shadows and no drop shadows) for both objects. Users showed a mean $GDisp_x$ of 36.39mm ± 53.57 in the no drop shadows condition across all positions, and a higher mean of 43.47mm ± 35.36 in the drop shadows condition. For spheres, users showed a mean $GDisp_x$ of 20.92mm ± 36.83 in the no drop shadows condition across all positions, and again a higher mean of 27.43mm ± 32.25 in the drop shadows condition.

Users showed higher $GDisp_x$ in the drop shadows condition, where drop shadows outperformed the no drop shadows condition in only 14 out of the 108 trials in this study with only 5 out of the 14 trials showing statistical significance (see Table 7.4 [page 155]). As shown by the clusters in Figures 7.6a [page 156] and 7.6c [page 156], adding drop shadows as a depth cue also increased the range of $GDisp_x$ for both objects, for cubes $GDisp_x$ ranged from 4.89mm ± 33.18 to 76.10mm ± 33.73 in the no drop shadows condition, and a higher range from 12.87mm ± 24.68 to 96.25mm ± 55.38 in the drop shadows condition. Similarly for spheres, $GDisp_x$ ranged in the no drop shadows conditions from 0.59mm ± 24.48 to 57.39mm ± 72.55, and from 0.75mm ± 13.99 to 74.32mm ± 26.65 in the drop shadows condition.

Users presented the lowest $GDisp_x$ in the 2000mm z plane for cubes in both conditions (drop shadows and no drop shadows), with a mean $GDisp_x$ of 36.88mm ± 28.11 in the drop shadows condition, and a mean of 27.40mm ± 66.60 in the no drop shadows condition across all positions. Similarly for spheres, users also presented the lowest $GDisp_x$ in the 2000mm z plane in

Table 7.4: Descriptive Statistics of $GDisp_x$ (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*)

2000mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	No Drop Shadows	30.38 ± 23.79*	33.33 ± 18.21	13.17 ± 21.38*
		Drop Shadows	41.60 ± 24.94*	32.68 ± 17.30	26.51 ± 21.82*
	Centre	No Drop Shadows	39.97 ± 26.28*	31.59 ± 31.41	-18.48 ± 163.64*
		Drop Shadows	48.55 ± 32.04*	30.97 ± 22.62	23.86 ± 21.97*
	Bottom	No Drop Shadows	54.21 ± 69.55*	36.19 ± 31.93*	26.22 ± 31.30*
		Drop Shadows	57.64 ± 37.82*	39.20 ± 27.63*	30.89 ± 23.45*
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	19.57 ± 18.97	14.25 ± 14.12*	0.59 ± 24.48*
		Drop Shadows	21.60 ± 17.35	21.13 ± 11.76*	9.56 ± 12.21*
	Centre	No Drop Shadows	26.05 ± 22.96*	17.73 ± 29.29	6.25 ± 16.73*
		Drop Shadows	23.69 ± 17.92*	19.72 ± 19.80	10.70 ± 11.02*
	Bottom	No Drop Shadows	37.38 ± 51.28*	26.13 ± 28.00	16.57 ± 37.96
		Drop Shadows	39.83 ± 30.89*	27.94 ± 32.09	10.02 ± 13.25
2200mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	51.74 ± 31.14	32.96 ± 17.90*	3.46 ± 31.25*
		Drop Shadows	49.12 ± 17.01	40.24 ± 20.42*	18.17 ± 21.54*
	Centre	No Drop Shadows	59.06 ± 34.43	43.79 ± 36.11*	17.27 ± 33.57*
		Drop Shadows	63.49 ± 27.69	49.58 ± 29.30*	22.31 ± 18.05*
	Bottom	No Drop Shadows	70.28 ± 51.83*	51.88 ± 41.57*	29.52 ± 41.97
		Drop Shadows	84.67 ± 37.50*	62.04 ± 29.76*	29.15 ± 23.17
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	42.16 ± 42.15*	28.94 ± 39.41*	0.72 ± 18.67
		Drop Shadows	35.74 ± 24.52*	26.15 ± 13.68*	0.75 ± 13.99
	Centre	No Drop Shadows	37.20 ± 36.80	22.80 ± 16.09	5.24 ± 15.18
		Drop Shadows	40.98 ± 26.75	23.56 ± 21.02	4.57 ± 16.37
	Bottom	No Drop Shadows	38.88 ± 37.74*	30.11 ± 24.95*	9.40 ± 22.42*
		Drop Shadows	67.96 ± 45.52*	41.68 ± 31.42*	17.43 ± 23.56*
2400mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	76.10 ± 33.73*	33.86 ± 25.01*	3.42 ± 46.12*
		Drop Shadows	57.41 ± 25.49*	40.83 ± 19.03*	20.83 ± 26.70*
	Centre	No Drop Shadows	73.46 ± 37.03*	45.18 ± 35.16	4.89 ± 33.18*
		Drop Shadows	83.61 ± 32.58*	43.93 ± 28.04	13.09 ± 27.28*
	Bottom	No Drop Shadows	80.80 ± 45.17*	45.76 ± 43.23*	12.47 ± 43.12
		Drop Shadows	96.25 ± 55.38*	54.23 ± 37.94*	12.87 ± 24.68
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	34.82 ± 51.37*	9.24 ± 25.66*	-0.88 ± 36.41
		Drop Shadows	59.83 ± 25.51*	23.56 ± 16.83*	4.33 ± 18.30
	Centre	No Drop Shadows	47.50 ± 30.33*	23.61 ± 23.22*	-12.27 ± 15.20
		Drop Shadows	74.32 ± 26.65*	35.98 ± 26.78*	-8.56 ± 25.59
	Bottom	No Drop Shadows	57.39 ± 72.55	33.55 ± 23.86	-8.06 ± 28.75*
		Drop Shadows	69.75 ± 38.47	34.08 ± 30.69	4.22 ± 23.09*



(a) Drop shadows - Cubes

(b) No drop shadows - Cubes

(c) Drop shadows - Spheres

(d) No drop shadows - Spheres

Figure 7.6: *gmp* placement in the x and y axes for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm). 7.6a and 7.6c: Drop shadows used. 7.6b and 7.6d: No Drop shadows used. Density heat maps indicate *gmp* placement across participants (red indicates higher density)

both conditions (drop shadows and no drop shadows) with a mean $GDisp_x$ of 20.47mm ± 21.97 across all positions in the drop shadows condition, and a mean $GDisp_x$ of 7.19mm ± 22.82 in the no drop shadows condition. This preference for the furthest z plane from users (2000mm) is potentially attributed to the spatial position of this particular plane being at the extremity of the average arm reach of users, thus users, in this case, were accurate in terms of grasp placement in the x axis as the environment design did not allow the arm reach to extend beyond this plane, thus reducing the amount of potential errors in grasp placement in the x axis.

As shown in Figures 7.3h [page 148] and 7.4h [page 149], lowest mean $GDisp_x$ was shown by users in the Top positions alongside the y axis across all z planes for cubes with a mean $GDisp_x$ of 36.38mm ± 25.03 in the drop shadows condition, and a mean $GDisp_x$ of 30.94mm ± 36.24 in the no drop shadows condition. Users also presented lowest $GDisp_x$ in Top positions for spheres with a mean $GDisp_x$ of 22.52mm ± 24.48 in the drop shadows condition, and a mean $GDisp_x$ of 16.60mm ± 35.69 in the no drop shadows condition.

Alongside the x axis (see Figures 7.3c [page 148] and 7.4c [page 149]), users showed the lowest $GDisp_x$ in Right positions across all z planes for cubes with a mean $GDisp_x$ of 21.97mm ± 24.10 in the drop shadows and a mean $GDisp_x$ of 10.22mm ± 65.61 in the no drop shadows condition. Right positions also provided the lowest $GDisp_x$ in grasping spheres with a mean $GDisp_x$ of 5.89mm ± 19.47 in the drop shadows condition, and a mean $GDisp_x$ of 1.95mm ± 26.64 in the no drop shadows condition.

Even though the addition of drop shadows as a depth cue did not reduce user $GDisp_x$ in the majority of positions in this study, it did reduce $GDisp_x$ variation in the majority of positions for both objects (see SD values in Table 7.4 [page 155]). Thus even though user $GDisp_x$ was generally higher in the drop shadows condition, grasp placement was more consistent in comparison to the no drop shadows condition across all users (see Figures 7.6a [page 156] and 7.6c [page 156]).

$GDisp_x$ is influenced by the gmp of a user in this study, where ideally gmp should be aligned with the omp in a perfected medium wrap grasp resulting in no $GDisp_x$. However, in reality, gmp is largely dependent on individual differences and how users interpret and perform a medium wrap grasp. For example if 10 users grasp a virtual object, it is highly likely that the grasps performed will have varying amounts of $GDisp_x$ that can accumulate to a high $GDisp_x$ mean across all users and positions due to individual differences, however these varying amounts of $GDisp_x$ are also most likely comparable and unnoticeable the users themselves, as gmp placement in the x axis is usually within the bounds of an object's size or in some cases penetrating the bounds of the virtual object.

***GDisp*$_y$**

$GDisp_y$ presented the lowest displacement in this study for both object and conditions (drop shadows and no drop shadows) under test. Users grasping cubes showed a mean $GDisp_y$ of 9.12mm ± 41.94 in the no drop shadows condition and a lower mean of 8.97mm ± 28.43 in the drop shadows condition. In contrast for spheres, a lower mean $GDisp_y$ of 5.22mm ± 47.49

Table 7.5: Descriptive Statistics of $GDisp_y$ (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*)

2000mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	No Drop Shadows	-12.23 ± 17.65*	-11.82 ± 36.64	-10.91 ± 19.41*
		Drop Shadows	-7.82 ± 16.45*	-10.15 ± 17.22	0.55 ± 11.11*
	Centre	No Drop Shadows	-3.48 ± 11.10*	5.03 ± 15.32	34.13 ± 100.20*
		Drop Shadows	1.44 ± 14.48*	1.91 ± 14.08	3.25 ± 13.61*
	Bottom	No Drop Shadows	24.81 ± 62.66	26.12 ± 25.17*	13.95 ± 23.12*
		Drop Shadows	13.58 ± 18.60	21.11 ± 21.47*	23.00 ± 23.99*
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	-12.56 ± 20.68	-7.50 ± 16.17*	-10.92 ± 29.53*
		Drop Shadows	-9.96 ± 14.71	-4.67 ± 9.50*	-3.50 ± 16.51*
	Centre	No Drop Shadows	-4.12 ± 14.83*	4.52 ± 48.29*	-1.21 ± 13.85*
		Drop Shadows	-0.63 ± 12.34*	1.69 ± 10.40*	2.79 ± 10.84*
	Bottom	No Drop Shadows	16.86 ± 31.56	20.28 ± 42.65*	19.13 ± 34.83
		Drop Shadows	16.24 ± 35.44	20.05 ± 25.02*	16.12 ± 21.71
2200mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	-28.59 ± 29.66*	-21.67 ± 22.23*	-8.30± 19.27*
		Drop Shadows	-12.36 ± 16.14*	-14.92 ± 10.52*	-0.85 ± 13.55*
	Centre	No Drop Shadows	12.33 ± 11.95	4.70 ± 11.42*	8.90 ± 17.63
		Drop Shadows	10.46 ± 10.17	11.83 ± 22.73*	8.03 ± 23.52
	Bottom	No Drop Shadows	27.92 ± 34.98	37.08 ± 35.89	31.28 ± 31.05
		Drop Shadows	29.50 ± 35.01	28.41 ± 29.83	28.22 ± 27.21
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	-33.72 ± 66.25*	-26.63 ± 64.09*	-8.66 ± 17.31
		Drop Shadows	-15.49 ± 18.53*	-11.98 ± 13.34*	-7.13 ± 17.35
	Centre	No Drop Shadows	1.73 ± 13.38	-0.13 ± 13.46	7.35 ± 14.01*
		Drop Shadows	1.12 ± 13.89	2.00 ± 14.34	6.98 ± 16.03*
	Bottom	No Drop Shadows	25.31 ± 30.29*	38.82 ± 43.26	25.34 ± 25.90
		Drop Shadows	35.44 ± 35.78*	31.93 ± 35.51	27.77 ± 23.26
2400mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	-42.99 ± 33.59*	-19.63 ± 26.17	0.34 ± 42.24
		Drop Shadows	-22.27 ± 16.18*	-23.52 ± 23.48	-4.63 ± 18.25
	Centre	No Drop Shadows	10.78 ± 16.92	6.85 ± 12.83	16.77 ± 18.30*
		Drop Shadows	10.01 ± 14.25	9.07 ± 20.51	9.22 ± 16.61*
	Bottom	No Drop Shadows	45.69 ± 39.88*	55.96 ± 39.59*	43.12 ± 36.22
		Drop Shadows	36.37 ± 32.58*	48.91 ± 36.80*	43.81 ± 28.64
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	-40.85 ± 32.43*	-26.55 ± 20.42*	-10.58 ± 23.59*
		Drop Shadows	-36.16 ± 18.74*	-23.53 ± 26.05*	-2.23 ± 16.00*
	Centre	No Drop Shadows	1.59 ± 17.13	2.52 ± 16.83*	6.83 ± 17.07
		Drop Shadows	2.01 ± 20.13	8.57 ± 22.22*	10.53 ± 15.20
	Bottom	No Drop Shadows	77.40 ± 132.78*	39.11 ± 37.17	37.60 ± 31.32*
		Drop Shadows	36.57 ± 37.46*	43.76 ± 42.15	42.75 ± 32.96*

was found in the no drop shadows condition and a higher mean of 7.07mm ± 30.38 in the drop shadows condition.

Similar to $GDisp_x$, users also showed higher mean $GDisp_y$ in the drop shadows condition, where the drop shadows condition outperformed the no drop shadows condition in just 38 out of the 108 trials in this study, with 24 out of the 38 showing statistical significance (see Table 7.5 [page 158]). However adding drop shadows reduced the range and variation of $GDisp_y$ for both objects, where mean $GDisp_y$ in grasping cubes ranged from 0.34mm ± 42.24 to 55.96mm ± 39.59 in the no drop shadows condition, and a narrower range from 0.55mm ± 11.11 to 43.81mm ± 28.64 in the drop shadows condition (see Figures 7.6a [page 156] and 7.6c [page 156]). Similarly for spheres, mean $GDisp_y$ ranged from -0.13mm ± 13.46 to 77.40mm ± 132.78 in the no drop shadows condition, and from -0.63mm ± 12.34 to 43.76mm ± 42.15 in the drop shadows condition. This shows that even though drop shadows did not reduce $GDisp_y$ in all trials in this study, it made grasp placement along the y axis more consistent, thus users showed more confidence in gmp placement across all positions in the y axis in the drop shadows condition in comparison to the no drop shadows condition.

Users presented lowest $GDisp_y$ in the 2000mm z plane for cubes in the drop shadows condition with a mean $GDisp_y$ of 5.21mm ± 20.43, and in the 2200mm z plane in the no drop shadows condition with a higher mean $GDisp_y$ of 7.07mm ± 33.48. For spheres, users presented lowest mean $GDisp_y$ in the 2000mm z plane for both conditions, with a mean of 4.24mm ± 21.64 in the drop shadows condition, and a lower $GDisp_y$ mean of 2.72mm ± 32.81 in the no drop shadows condition. Similar to $GDisp_x$, this plane shows lowest $GDisp_y$ due to its spatial position at the extremity of the arm reach, thus limiting any potential errors by users along the y axis. As shown in Figures 7.3i [page 148] and 7.4i [page 149], lowest mean $GDisp_y$ was shown across all z planes by users in Centre positions for cubes in both conditions (drop shadows and no drop shadows) with a mean of 7.24mm ± 17.60 in the drop shadows condition, and a higher mean of 10.67mm ± 37.45 in the no drop shadows condition. Centre positions also showed lowest $GDisp_y$ for spheres in both conditions with a mean of 3.89mm ± 15.91 in the drop shadows condition and a lower mean of 2.12mm ± 21.79 in the no drop shadows condition. This preference for Centre and Right positions is expected and in alignment with findings of $GDisp_y$ as all users in this study were right-handed.

However alongside the x axis (see Figures 7.3d [page 148] and 7.4d [page 149]), users showed the lowest mean $GDisp_y$ in Left positions for grasping cubes in both conditions (drop shadows and no drop shadows), with a mean $GDisp_y$ of 6.55mm ± 27.59 in the drop shadows condition, and a mean $GDisp_y$ of 3.80mm ± 42.22 in the no drop shadows condition. Users were also more accurate in Left positions in grasping spheres in both the drop shadows and no drop shadows conditions, with a mean $GDisp_y$ of 3.24mm ± 33.35 in the drop shadows condition, and a mean $GDisp_y$ of 3.52mm ± 63.12 in the no drop shadows condition. This is surprising as all users in this study are right handed, however, this highlights how highly varied $GDisp_y$ can be between individual users, as Right positions were consistently second best in terms of low $GDisp_y$.

Similar to $GDisp_x$, $GDisp_y$ is also highly influenced by gmp placement that can differ from one user to the other due to individual differences in grasp performance, and even though adding drop shadows did not reduce $GDisp_y$ across all positions, it still made gmp placement along the y axis more consistent. In addition, $GDisp_y$ showed the lowest displacement in this study (see Table 7.4 [page 155]), this shows that gmp placement along the y axis was not problematic for individual users, as similar to gmp on the x axis ($GDisp_x$) individual users provide comparable results along the y axis ($GDisp_y$).

***GDisp*$_z$**

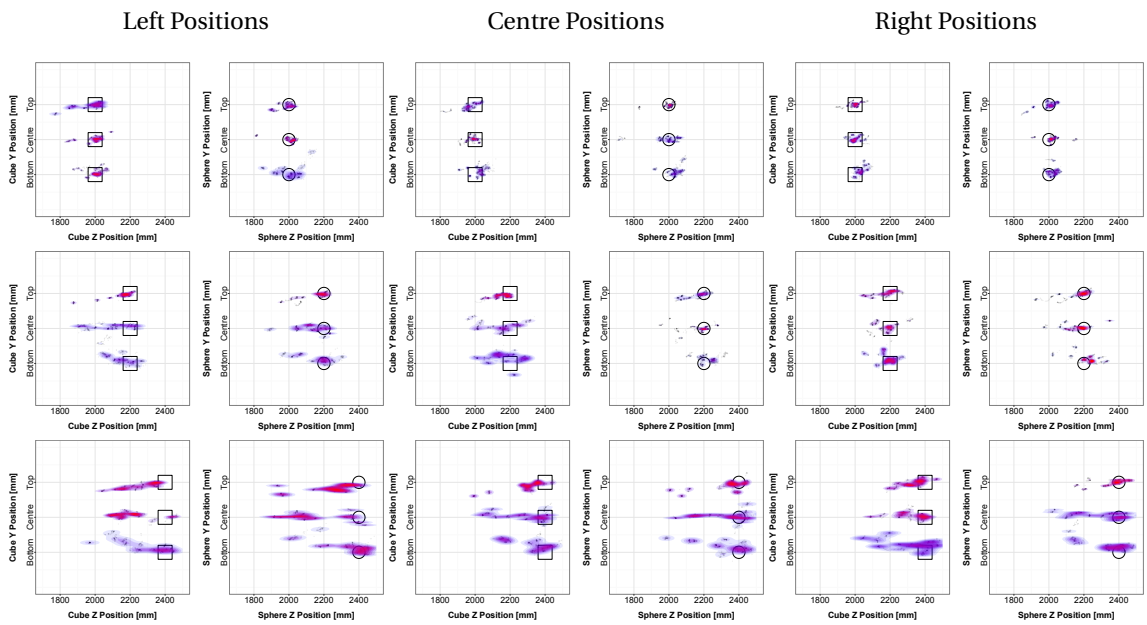
$GDisp_z$ presented the highest displacement in this study in both conditions (drop shadows and no drop shadows) for both objects (see Figures 7.7a [page 162] and 7.7b [page 162]). For cubes, users showed a mean $GDisp_z$ of -56.80mm $\pm$ 98.59 in the drop shadows condition across all positions, and a higher mean $GDisp_z$ of -80.20mm $\pm$ 155.06 in the no drop shadows condition. Similarly for spheres, a mean $GDisp_z$ of -46.54mm $\pm$ 111.52 was found in the drop shadows condition, and a higher mean $GDisp_z$ of -57.66mm $\pm$ 136.75 in the no drop shadows condition. Users showed a high variation in estimating the z position of virtual objects across in the majority of positions in this study (see Table 7.6 [page 161]). This shows that depth estimation in exocentric AR is problematic, and this is evident by the wide ranges of mean $GDisp_z$ in both conditions (drop shadows and no drop shadows), where for cubes it ranged from 1.54mm $\pm$ 44.97 to -183.51mm $\pm$ 108.47 in the drop shadows condition, and from 5.30mm $\pm$ 46.58 to -268.27mm $\pm$ 216.36 in the no drop shadows condition. For spheres, mean $GDisp_z$ ranged from -0.58mm $\pm$ 52.39 to -224.16mm $\pm$ 167.17 in the drop shadows condition, and a narrower range in the no drop shadows condition from -2.36mm $\pm$ 62.61 to -166.79mm $\pm$ 187.07. Users also showed a high tendency of underestimating the depth of the virtual objects, where 91 out of the 108 trials in this study were negative regardless of the drop shadows used (see Table 7.6 [page 161]), this shows that users placed their gmp at a closer depth in relation to the sensor than the depth of the virtual object.

Adding drop shadows as a depth cue alongside occlusion generally reduced the mean $GDisp_z$ across all positions, where drop shadows outperformed the no drop shadows condition in the majority of trials (35 out of the 54 individual trials in this test), with 27 of the 35 showing statistical significance ($p < 0.01$). Adding drop shadows also reduced the amount of variation of users in depth estimation, thus even if users were not accurate in depth estimation, they were more confident in their gmp placement along the z axis and were more clustered around the object as seen in Figure 7.7a [page 162].

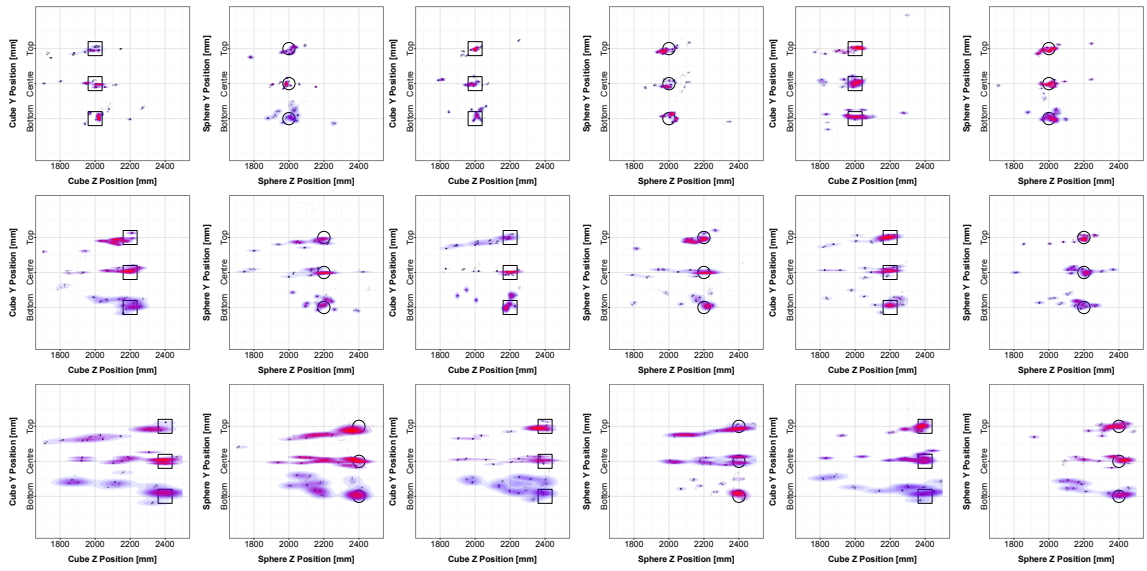
Users showed lowest mean $GDisp_z$ in the 2000mm z plane for grasping cubes in both conditions (drop shadows and no drop shadows), where users showed a mean $GDisp_z$ of -7.10mm $\pm$ 56.91 in the drop shadows condition, and a mean $GDisp_z$ of -17.09mm $\pm$ 100.62 in the no drop shadows condition. Similarly for grasping spheres, users showed a mean $GDisp_z$ of 4.11mm $\pm$ 54.35 in the drop shadows condition, and a mean $GDisp_z$ of -2.05mm $\pm$ 91.69 in the no drop shadows condition. This preference for the furthest z plane from users (2000mm) can again be

Table 7.6: Descriptive Statistics of $GDisp_z$ (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*)

2000mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cube (Constant Size - 80mm)	Top	No Drop Shadows	-46.58 ± 104.36*	-12.92 ± 106.71*	-66.34 ± 115.41*
		Drop Shadows	-38.79 ± 77.79*	-30.19 ± 46.90*	-4.05 ± 32.88*
	Centre	No Drop Shadows	-24.81 ± 97.87	-46.51 ± 79.18*	-10.23 ± 107.75*
		Drop Shadows	-31.07 ± 95.01	-15.35 ± 29.15*	2.80 ± 28.89*
	Bottom	No Drop Shadows	42.80 ± 83.21*	5.30 ± 46.58*	5.52 ± 101.00*
		Drop Shadows	1.54 ± 44.97*	18.90± 42.76*	32.27 ± 26.43*
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	-18.04 ± 84.73*	-7.24 ± 45.40*	-22.72 ± 144.97*
		Drop Shadows	-19.74 ± 42.74*	-7.40 ± 45.76*	-3.01 ± 47.42*
	Centre	No Drop Shadows	-32.91 ± 102.52*	-2.36 ± 62.61	4.64 ± 103.93*
		Drop Shadows	-0.58 ± 52.39*	-13.50 ± 78.29	20.49 ± 45.69*
	Bottom	No Drop Shadows	12.54 ± 81.24	38.38 ± 87.29*	9.27 ± 48.81*
		Drop Shadows	3.48 ± 68.89	34.44 ± 39.69*	22.81 ± 23.26*
2200mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	-139.70 ± 158.97*	-128.89 ± 121.07*	-89.84 ± 121.12*
		Drop Shadows	-60.44 ± 86.34*	-54.49 ± 47.15*	-33.96 ± 62.34*
	Centre	No Drop Shadows	-63.06 ± 112.05*	-70.29 ± 120.38	-71.31 ± 118.05*
		Drop Shadows	-105.92 ± 111.57*	-58.08 ± 96.21	-21.53 ± 55.90*
	Bottom	No Drop Shadows	-58.91 ± 117.23*	-26.84 ± 144.95	-42.36 ± 116.05
		Drop Shadows	-51.68 ± 68.51*	-28.44 ± 82.98	-17.58 ± 48.10
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	-121.74 ± 176.04*	-37.12 ± 137.20	-40.15 ± 103.35*
		Drop Shadows	-62.72 ± 81.86*	-42.65 ± 58.03	-53.15 ± 87.94*
	Centre	No Drop Shadows	-83.72 ± 165.46	-95.75 ± 139.89*	-25.05 ± 117.84*
		Drop Shadows	-73.22 ± 88.64	-25.76 ± 73.24*	-44.94 ± 81.75*
	Bottom	No Drop Shadows	-10.37 ± 81.18*	-42.35 ± 124.17*	-20.35 ± 140.38*
		Drop Shadows	-39.07 ± 89.23*	11.01 ± 61.87*	16.62 ± 73.08*
2400mm Z Plane					
Object Type	Position (y)	Rendering	Position (x)		
			Left	Centre	Right
Cubes (Constant Size - 80mm)	Top	No Drop Shadows	-268.27 ± 216.36*	-119.59 ± 162.43*	-60.42 ± 137.05*
		Drop Shadows	-153.22 ± 97.46*	-106.55 ± 111.32*	-68.85 ± 69.94*
	Centre	No Drop Shadows	-182.08 ± 219.65*	-147.90 ± 165.50	-101.71 ± 160.32*
		Drop Shadows	-183.51 ± 108.47*	-130.39 ± 129.38	-89.17 ± 93.47*
	Bottom	No Drop Shadows	-163.35 ± 214.82	-143.42 ± 140.32*	-133.56 ± 205.61
		Drop Shadows	-131.96 ± 137.85	-78.80 ± 102.29*	-95.07 ± 130.12
Spheres (Constant Size - 70mm)	Top	No Drop Shadows	-166.79 ± 187.07*	-126.73 ± 147.45*	-63.32 ± 130.68
		Drop Shadows	-177.49 ± 148.40*	-82.23 ± 133.41*	-45.94 ± 88.63
	Centre	No Drop Shadows	-154.90 ± 143.01*	-131.93 ± 146.31*	-68.31 ± 114.48*
		Drop Shadows	-224.16 ± 167.17*	-86.75 ± 133.05*	-113.32 ± 132.93*
	Bottom	No Drop Shadows	-165.13 ± 138.32*	-70.64 ± 111.69	-113.98 ± 141.19*
		Drop Shadows	-90.40 ± 143.18*	-84.96 ± 132.45	-74.41 ± 116.30*



(a) *gmp* placement in the z axis for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm) in the drop shadows condition



(b) *gmp* placement in the z axis for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm) in the no drop shadows condition

Figure 7.7: *gmp* placement in the z axis for cubes (black squares) and spheres (black circles) of all participants in 3 z planes (starting from top row in the order: 1400mm - 1600mm - 1800mm) using 7.7a: drop shadows and 7.7b: No drop shadows. Density heat maps indicate *gmp* placement across participants (red indicates higher density)

attributed to the spatial position of the plane, that limits the amount of false depth estimation due to its position at the extremity of the average arm reach of users. As shown in Figures 7.3j [page 148] and 7.4j [page 149], users presented the lowest mean $GDisp_z$ in the Bottom positions alongside the y axis across all z planes in grasping cubes in both conditions (drop shadows and no drop shadows), with a mean $GDisp_z$ of -38.98mm $\pm$ 99.40 in the drop shadows condition, and a mean $GDisp_z$ of -57.20mm $\pm$ 155.95 in the no drop shadows condition. Bottom positions also showed lowest mean $GDisp_z$ for spheres in both the drop shadows and no drop shadows conditions, with a mean of -22.28mm $\pm$ 103.04 in the drop shadows condition, and a mean $GDisp_z$ of -40.29mm $\pm$ 126.65 in the no drop shadows condition. This is surprising as the Bottom positions consistently presented the highest user errors in terms of GAp, $GDisp_x$, $GDisp_y$ and also presented the longest completion times for both conditions and object types. This can potentially be attributed to the position of the arm during grasping, wherein Bottom positions the position of the arm is not obstructing the view of the user of the feedback monitor, this may have lead to more accurate depth estimation due to an unobstructed judgement of the position of the hand and fingers, more so than Top and Centre positions where the arm can obstruct the view of the feedback during grasping.

Alongside the x axis (see Figures 7.3e [page 148] and 7.4e [page 149]), users showed the lowest mean $GDisp_z$ in Right positions across all z planes for grasping cubes in both conditions (drop shadows and no drop shadows), with a mean $GDisp_z$ of -32.79mm $\pm$ 79.92 in the drop shadows condition and a mean $GDisp_z$ of -63.36mm $\pm$ 140.99 in the no drop shadows condition. Users also showed lowest mean $GDisp_z$ in Right positions for grasping spheres in both conditions (drop shadows and no drop shadows), with a mean $GDisp_z$ of -30.54mm $\pm$ 95.31 in the drop shadows condition, and a mean $GDisp_z$ of -37.77mm $\pm$ 124.95 in the no drop shadows condition.

7.3.7 Results - Object Type

2000mm Z plane

In the drop shadows condition, a statistically significant difference in GAp between cubes and spheres in different positions was found (Z = 2.00 $\times 10^7$, p < 0.01) No statistically significant difference in completion time between cubes and spheres in different positions was found (Z = 1.47 $\times 10^7$, p > 0.01). A statistically significant difference in $GDisp$ between cubes and spheres in different positions was also found in $GDisp_x$ (Z = 2.07 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.52 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.25 $\times 10^7$, p < 0.01).

In the no drop shadows condition, a statistically significant difference in GAp between cubes and spheres in different positions was found (Z = 1.75 $\times 10^7$, p < 0.01). A statistically significant difference in completion time between cubes and spheres in different positions was also found (Z = 1.65 $\times 10^7$, p < 0.01). A statistically significant difference in $GDisp$ between cubes and spheres in different positions was also found in $GDisp_x$ (Z = 1.90 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.61 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.33 $\times 10^7$, p < 0.01).

In between the two conditions (drop shadows and no drop shadows), grasping cubes showed statistical significant differences between the drop shadows and no drop shadows conditions in GAp (Z = 1.41 $\times 10^7$, p < 0.01), $GDisp_x$ (Z = 1.26 $\times 10^7$, p < 0.01), $GDisp_z$ (Z = 1.35 $\times 10^7$, p < 0.01) and completion time (Z = 1.72 $\times 10^7$, p < 0.01). No statistical significant difference was found in $GDisp_y$ between the drop shadows and no drop shadows conditions (Z = 1.42 $\times 10^7$, p > 0.01). Spheres also showed statistical significant differences between the drop shadows and no drop shadows conditions in GAp (Z = 1.71 $\times 10^7$, p < 0.01), $GDisp_x$ (Z = 1.33 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.30 $\times 10^7$, p < 0.01), $GDisp_z$ (Z = 1.30 $\times 10^7$, p < 0.01) and completion time (Z = 1.55 $\times 10^7$, p < 0.01).

2200mm Z plane

In the drop shadows condition, a statistically significant difference in GAp between cubes and spheres in different positions was found (Z = 1.82 $\times 10^7$, p < 0.01). No statistically significant difference in completion time between cubes and spheres in different positions was found (Z = 1.48 $\times 10^7$, p < 0.01). A statistically significant difference in $GDisp$ between cubes and spheres in different positions was also found in $GDisp_x$ (Z = 1.99 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.55 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.27 $\times 10^7$, p < 0.01).

In the no drop shadows condition, a statistically significant difference in GAp between cubes and spheres in different positions was found (Z = 1.78 $\times 10^7$, p < 0.01). A statistically significant difference in completion time between cubes and spheres in different positions was also found (Z = 1.76 $\times 10^7$, p < 0.01). A statistically significant difference in $GDisp$ between cubes and spheres in different positions was also found in $GDisp_x$ (Z = 1.85 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.54 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.32 $\times 10^7$, p < 0.01).

In between the two conditions (drop shadows and no drop shadows), grasping cubes showed statistical significant differences between the drop shadows and no drop shadows conditions in GAp (Z = 1.53 $\times 10^7$, p < 0.01), $GDisp_x$ (Z = 1.29 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.39 $\times 10^7$, p < 0.01), $GDisp_z$ (Z = 1.38 $\times 10^7$, p < 0.01) and completion time (Z = 1.63 $\times 10^7$, p < 0.01). Spheres also showed statistical significant differences between the drop shadows and no drop shadows conditions in GAp (Z = 1.60 $\times 10^7$, p < 0.01), $GDisp_x$ (Z = 1.35 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.38 $\times 10^7$, p < 0.01), $GDisp_z$ (Z = 1.38 $\times 10^7$, p < 0.01) and completion time (Z = 1.37 $\times 10^7$, p < 0.01).

2400mm Z plane

In the drop shadows condition, a statistically significant difference in GAp between cubes and spheres in different positions was found (Z = 1.78 $\times 10^7$, p < 0.01) No statistically significant difference in completion time between cubes and spheres in different positions was found (Z = 1.46 $\times 10^7$, p > 0.01). A statistically significant difference in $GDisp$ between cubes and spheres in different positions was also found in $GDisp_x$ (Z = 1.75 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.52 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.32 $\times 10^7$, p < 0.01).

In the no drop shadows condition, a statistically significant difference in GAp between cubes and spheres in different positions was found (Z = 1.79 $\times 10^7$, p < 0.01). A statistically significant difference in completion time between cubes and spheres in different positions was also found

(Z = 1.84 $\times 10^7$, p < 0.01). A statistically significant difference in $GDisp$ between cubes and spheres in different positions was also found in $GDisp_x$ (Z = 1.85 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.60 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.35 $\times 10^7$, p < 0.01).

In between the two conditions (drop shadows and no drop shadows), grasping cubes showed statistical significant differences between the drop shadows and no drop shadows conditions in GAp (Z = 1.40 $\times 10^7$, p < 0.01), $GDisp_x$ (Z = 1.38 $\times 10^7$, p < 0.01) and completion time (Z = 1.68 $\times 10^7$, p < 0.01). No statistical significant difference between the drop shadows and no drop shadows conditions was found in $GDisp_y$ (Z = 1.50 $\times 10^7$, p < 0.01) and $GDisp_z$ (Z = 1.47 $\times 10^7$, p < 0.01). Spheres also showed statistical significant differences between the drop shadows and no drop shadows conditions in $GDisp_x$ (Z = 1.25 $\times 10^7$, p < 0.01), $GDisp_y$ (Z = 1.39 $\times 10^7$, p < 0.01), $GDisp_z$ (Z = 1.40 $\times 10^7$, p < 0.01) and completion time (Z = 1.29 $\times 10^7$, p < 0.01). No statistical significant difference between the drop shadows and no drop shadows conditions was found in GAp (Z = 1.43 $\times 10^7$, p < 0.01).

7.3.8 Analysis - Object Type

In this section, findings for different object types (cubes and spheres) are reported per each z plane, and not for each individual position in this study to avoid repetition with results previously reported.

In GAp users consistently overestimated object size in both conditions and object types, this can again be attributed to the lack of tactile feedback in freehand grasping. In terms of accuracy, users showed more accuracy in matching their GAp to object size in grasping cubes than spheres, this was consistent for all conditions (drop shadows and no drop shadows) and z planes with the exception of the 2000mm z plane where users performed more accurately in grasping spheres than cubes in the drop shadows condition (see Table 7.7 [page 166]). However, users showed less variation and more confidence in grasping spheres than cubes across all conditions and planes, with the exception of the 2400mm z plane where users showed less variation for grasping cubes in the no drop shadows condition (see SD values in Table 7.7 [page 166]). Adding drop shadows as a depth cue generally resulted in lower accuracy in GAp matching to objects size for both objects, with the exception of the 2200mm z plane where adding drop shadows improved user performance in matching GAp to object size for both cubes and spheres across all positions, and the 2000mm z plane where adding drop shadows improved matching GAp to object size for spheres, but not cubes.

In $GDisp_x$, users consistently showed lower mean $GDisp_x$ and variation for spheres than cubes across all planes and conditions (drop shadows and no drop shadows), this is potentially attributed to the size of spheres being smaller than cubes, thus the chance of error in $GDisp_x$ estimation is naturally smaller for spheres than the bigger cubes (see Table 7.7 [page 166]). In addition, the nature of spherical shapes consisting of one surface also made it easier for users to align their gmp to the omp naturally without being constrained by visible faces that may in-

Table 7.7: Descriptive Statistics of Object Type (Mean ± SD). Significant differences between the drop shadows and no drop shadows conditions ($p < 0.01$) are marked with (*)

2000mm z Plane						
Object Type	**Rendering**	GAp **[mm]**	$GDisp_{\mathbf{x}}$ **[mm]**	$GDisp_{\mathbf{y}}$ **[mm]**	$GDisp_{\mathbf{z}}$ **[mm]**	**Time[s]**
Cubes	No Drop Shadows	5.93 ± 21.06 *	27.40 ± 66.60 *	7.29 ± 47.28	-17.09 ± 100.62*	10.02 ± 5.73*
(Size - 80mm)	Drop Shadows	7.38 ± 21.28 *	36.88 ± 28.11*	5.21 ± 20.43	-7.10 ± 56.91*	8.25 ± 4.01*
Spheres	No Drop Shadows	10.33 ± 20.08*	18.28 ± 30.97*	2.72 ± 32.81*	-2.05 ± 91.69*	8.90 ± 5.28*
(Size - 70mm)	Drop Shadows	4.62 ± 18.78*	20.47 ± 21.97*	4.24 ± 21.64*	4.11 ± 54.35*	7.96 ± 3.25*
2200mm z Plane						
Object Type	**Rendering**	GAp **[mm]**	$GDisp_{\mathbf{x}}$ **[mm]**	$GDisp_{\mathbf{y}}$ **[mm]**	$GDisp_{\mathbf{z}}$ **[mm]**	**Time [s]**
Cubes	No Drop Shadows	4.77 ± 22.25*	40.00 ± 41.67*	7.07 ± 33.48*	-76.80 ± 131.10*	9.01 ± 4.30*
(Size - 80mm)	Drop Shadows	2.26 ± 24.43*	46.53 ± 32.79*	9.81 ± 27.64*	-48.01 ± 80.33*	7.95 ± 3.82*
Spheres	No Drop Shadows	7.37 ± 21.67*	23.94 ± 33.26*	3.27 ± 44.05*	-52.95 ± 139.24*	7.70 ± 3.96*
(Size - 70mm)	Drop Shadows	3.67 ± 19.95*	28.76 ± 32.40*	7.85 ± 28.93*	-34.87 ± 83.24*	7.62 ± 2.96*
2400mm z Plane						
Object Type	**Rendering**	GAp **[mm]**	$GDisp_{\mathbf{x}}$ **[mm]**	$GDisp_{\mathbf{y}}$ **[mm]**	$\boldsymbol{GDisp}_z$ [mm]	**Time [s]**
Cubes	No Drop Shadows	4.86 ± 22.43*	41.77 ± 48.11*	12.99 ± 43.56	-146.70 ± 190.81	10.08 ± 5.44*
(Size - 80mm)	Drop Shadows	5.33 ± 24.28*	47.01 ± 42.68*	11.89 ± 34.92	-115.28 ± 116.20	9.10 ± 5.80*
Spheres	No Drop Shadows	7.19 ± 23.82	20.55 ± 44.60*	9.67 ± 61.00*	-118.00 ± 146.82*	8.02 ± 4.93*
(Size - 70mm)	Drop Shadows	8.39 ± 24.69	33.06 ± 38.80*	9.14 ± 38.09*	-108.85 ± 144.38*	8.55 ± 4.11*

fluence the grasp structure and spatial position resulting in higher $GDisp_{\text{x}}$. This is particularly true for cubes where the faces, even if virtual, can hinder grasp placement in order to perform a grasp that is adjusted in accordance with the visible faces of the cube. For example, users grasping spheres in this study could cover the whole object with their grasp in the first attempt without having to reconstruct the posture of their grasp due to edges afterwards, whereas for cubes users naturally attempt to perform a grasp that is not only accurate, but is also naturally compliant with the geometrical features of the cube. Adding drop shadows as a depth cue reduced variation in gmp placement along the x axis for both objects (i.e. range of $GDisp_{\text{x}}$), thus users were more confident in their grasp placement in the drop shadows condition. However adding drop shadows increased the mean $GDisp_{\text{x}}$ for cubes and spheres, thus user performance in gmp placement along the x axis was not significantly improved by the addition of drop shadows, this can potentially be attributed to users being more focused on using the drop shadows depth cue to locate the object successfully in the shortest time possible rather than accurately grasp it.

In $GDisp_{\text{y}}$, users again consistently showed lower mean $GDisp_{\text{y}}$ for spheres than cubes in all conditions (drop shadows and no drop shadows) and z planes (see Table 7.7 [page 166]), similar to $GDisp_{\text{x}}$ this can again be attributed to the smaller size and shape attributes of spheres. Adding drop shadows as a depth cue had a varying impact on $GDisp_{\text{y}}$ for cubes and sphere, where drop shadows reduced $GDisp_{\text{y}}$ for grasping cubes in two out of three z planes in this study (2000mm and 2400mm) across all positions, and in one z plane for spheres (2400mm). Drop shadows also reduced variation in grasping for both objects in all three z planes (see SD values in Table 7.7 [page 166]), thus users were more confident in their gmp placement along the y axis, however this did not necessarily lead to a significant improvement in $GDisp_{\text{y}}$.

In $GDisp_{\text{z}}$, users consistently underestimated the position of both objects (cubes and spheres) in all z planes and conditions (drop shadows and no drop shadows) by placing their gmp along the z axis in a depth that is closer to the sensor than the depth of the virtual object, with the exception of spheres in the 2000mm z plane for the drop shadows where users overestimated the z position. Users showed lower mean $GDisp_{\text{z}}$ in grasping spheres than cubes in all z planes and conditions (drop shadows and no drop shadows), thus users were more accurate in estimating the z position of spheres than cubes. This can again be attributed to the size and shape attributes of spheres. Adding drop shadows reduced $GDisp_{\text{z}}$ for both objects in all z planes, with the exception of the 2000mm z plane for spheres where drop shadows increased the mean $GDisp_{\text{z}}$ (see Table 7.7 [page 166]). Drop shadows also reduced variation in depth estimation for both objects in all z planes, thus users were more confident and more accurate in their gmp placement along the z axis with the presence of drop shadows. This shows that drop shadows can aid in depth perception of virtual objects when integrated in an exocentric AR environment as in this study, especially in the z axis where the real depth of objects cannot be perceived using only the feedback monitor and relying on occlusion alone as a depth cue can still result in false depth perception as shown as in Chapter 5 [page 66].

In completion time, users showed less task completion time in grasping spheres than cubes and this was consistent in all z planes (see Table 7.7 [page 166]). This shows that users required less time to locate spheres, this was evident during the study as users informally expressed that spheres are easier to interact with than cubes. The lesser need to adjust the posture of a grasp after locating spheres due to their smaller size and shape attributes could have also influenced completion time spent by users when grasping spheres. Adding drop shadows reduced completion time for both objects in all z planes, with the exception of the 2400mm z plane where users performed faster in the no drop shadows condition for spheres. Drop shadows also reduced variation in completion time for both objects in all z planes (see SD values in Table 7.7 [page 166]), thus users were more confident in the presence of drop shadows of making a decision that they were satisfied with their grasp. This shows that adding drop shadows can significantly reduce task completion time in freehand grasping by easing the process of locating a virtual object and allowing users in exocentric AR to realise the depth of an object by looking to its corresponding shadow on the floor.

7.3.9 Usability Analysis

SUS average score for the use of drop shadows in this study was 81.16 (SD = 11.56). This rating of drop shadows was higher than the rating of the no drop shadows condition (78.17 (SD = 14.13)) and is classified as "GOOD and highly acceptable" (Bangor et al., 2009*a*). This rating is also higher than the rating found for the dual view visual feedback method presented in Chapter 6 (a score of 64.50 (SD = 13.43) that is classified as "OK and marginally acceptable"). This shows that users preferred drop shadows more than dual view visual feedback as an additional depth cue, even though dual view visual feedback was more effective in terms of improving grasp accuracy and placement. Interestingly the no drop shadows condition, that is essentially the same method presented in Chapter 5 with occlusion being the only depth cue, also had a higher usability rating than the dual view visual feedback method. This illustrates that enabling users to be highly accurate in grasping performance (i.e. dual view visual feedback) may very well improve performance, but can significantly hinder usability of the system developed. User comments using post test questionnaires below provide general subjective insights regarding their experience in grasping virtual objects using dual view visual feedback, however these insights may not be directly representative of user performance and accuracy during interaction as these subjective responses were not measured against performance in this work.

In terms of easiness of tasks, the no drop shadows condition scored marginally higher (3.93 / 5.00) than the drop shadows condition (3.80 / 5.00). Interestingly users spent lower task completion times in the drop shadows condition, this is potentially due to the fact that grasping without drop shadows in this study is not challenged or corrected by any additional depth cues (i.e. drop shadows), this potentially leads users to perceive the task to be easier even if not accurate. In contrast with the addition of drop shadows interaction can be corrected using this additional depth cue, along with the additional cognitive load that is present with drop shad-

ows.

Users have also indicated that changes in virtual object position in between different z planes and in the x and y axes were more perceptible with the use of drop shadows (4.80 / 5.00) than without them (3.87 / 5.00). In addition, 9 out of 15 users (60.0%) have indicated that they relied on both drop shadows and occlusion that is implemented in the baseline setup for this work. This is in alignment with current research that states that depth cues can be more effective when used alongside shadows.

Hypothesis - Revisited

$H_{3.1}$: adding drop shadows in freehand grasping of virtual objects that change in position has no effect on a) grasp aperture and b) grasp displacement: **Rejected** as statistically significant results were found for drop shadows condition showing that adding drop shadows in grasping virtual objects that change in position has a significant effect on GAp, $GDisp$ in all axes (x, y and z).

$H_{3.2}$: Adding drop shadows in freehand grasping of virtual objects that change in position has no effect on task completion time: **Rejected** as statistically significant results were found for drop shadows condition showing that adding drop shadows in grasping virtual objects that change in position has a significant effect on task completion time.

7.4 Conclusions

This chapter presented a first study looking into the use of drop shadows to assist in freehand grasping of virtual objects in exocentric AR. The impact of drop shadows on grasp accuracy was quantified in two user experiments under two primary conditions: drop shadows and no drop shadows, where grasp accuracy in this study was measured using the proposed metrics in Chapter 4; GAp and $GDisp$ for grasping virtual objects that change position in the x, y and z axes. This study has also addressed the key problems found in Study 2 (Chapter 6) where dual view visual feedback was used, namely long task completion times and inaccurate estimation of virtual object size using GAp.

Findings in this study have illustrated that using drop shadows significantly reduces task completion times, that was one of the key problems found in Study 2 (Chapter 6). Drop shadows have also significantly improved depth estimation of virtual objects in the z axis (i.e. significantly reduced $GDisp_z$). Users also showed less variation in their grasp placement along the z axis under the drop shadows condition (i.e. lower $GDisp_z$ range across all users). These significant improvements in task completion times and depth estimation can be attributed to the additional visual information that drop shadows provide to users regarding positions of virtual objects, where they were able to locate virtual objects in different z planes even prior to the start of their grasping movements.

For the x and y axes (i.e. $GDisp_x$, $GDisp_y$), user performance was found to be comparable between the two conditions in this study (drop shadows and no drop shadows). This shows that using drop shadows does not necessarily improve grasp placement in relation to the virtual object's position in the x and y axes, as similarities in user grasp placement were found even with the addition of drop shadows. This is potentially due to the feedback method used being single view visual feedback, and given that this study reverted back to this baseline visual feedback method, problems in grasp placement that were found in Study 1 (Chapter 5) have re-emerged in this study.

Furthermore, this study also showed that users were more accurate in matching GAp to object size in the no drop shadows condition than the drop shadows condition. This is in alignment with the findings illustrated in Chapter 6 using dual view visual feedback and is attributed to the users focusing more on accurate grasp placement using the additional spatial cue provided to them (i.e. drop shadows or a secondary view) rather than size estimation. This again shows that size estimation using GAp is still problematic in exocentric AR, and users mainly show the highest accuracy in size estimation when they are not required to focus on any secondary spatial or depth cues in the environment.

Finally, the usability analysis of drop shadows using the SUS have shown that the use of drop shadows for freehand grasping was rated as good and highly acceptable. This rating was higher than that found for the no drop shadows condition. In addition, users also indicated that drop shadows made position changes of virtual objects in all axes more perceptible during freehand grasping. This is particularly important for exocentric AR systems where the user is not co-located with the virtual objects presented, and can potentially improve grasping performance by making users more aware of changes to the virtual information presented especially along the z axis where users are normally not able to visualise their interaction using single view visual feedback in an exocentric setting. Both conditions in this study (drop shadows and no drop shadows) were also rated higher in terms of usability than the dual view visual feedback method presented in Chapter 6. Thus even though dual view visual feedback was more effective in terms of improvements in depth estimation along the z axis (i.e. more accurate), users still preferred using drop shadows as an additional depth cue even if they were less accurate in depth estimation. This further emphasises the importance of considering the speed / accuracy trade-off that is associated with providing users with additional depth and visual cues during freehand grasping.

In conclusion, this chapter measured the usability and impact of drop shadows on the accuracy of freehand grasping of virtual objects that change in position and type in exocentric AR. Findings in this study showed that the use of drop shadows for freehand grasping is highly usable, and significantly improves task completion time, thus mitigating one of the key problems found in Chapter 6. Drop shadows have also significantly improved user depth estimation along the z axis. However, this study has also shown that using drop shadows does not improve size estimation using GAp that remains to be problematic in all the studies covered so far (1 to 3). In ad-

dition, the use of drop shadows in AR applications is not always feasible due to the limited FOV that current AR systems have (Grinshpoon et al., 2018), thus users are not always guaranteed to be able to visualise the whole environment with drop shadows as implemented in this study making it challenging to translate these methods to a wearable based setting in AR. In Chapter 8 these two problems will be addressed using user-based grasp tolerances that are based on the user errors found in Study 1 (Chapter 5). Grasp tolerances will aim to offer an alternative solution to drop shadows and dual view visual feedback, and to mitigate the lasting problem of virtual object size estimation using GAp. The next chapter will revisit the problems presented in this study by quantifying the impact of user-based grasp tolerances on freehand grasping performance in terms of task completion times. In addition, the usability of user-based grasp tolerances for freehand grasping will be discussed.

Chapter 8

Study 4: User Based Tolerances for Freehand Grasping

This work was published in the proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI) as "Freehand grasping in mixed reality: analysing variation during transition phase of interaction" (Al-Kalbani et al. (2017))

8.1 Introduction

This chapter will present the final user study in this work (see Figure 4.11 [page 64]) to assess the usability and impact of freehand grasping in exocentric AR. Findings in Chapter 7 have shown that freehand grasping performance can be improved using drop shadows, through significantly improving task completion time and usability. However implementing drop shadows in AR and AR applications is not always valid or feasible due to the limited FOV in current state of the art devices that mediate interaction between the human hand and virtual objects (Ren et al., 2016). In addition, the use of drop shadows also highlighted inaccurate size estimation as a problem in freehand grasping, this problem was present across all the studies in this thesis. It was evident in this work that accurate size estimation is problematic in freehand grasping, this is mainly due to the lack of tactile feedback on the hand and physical object features in freehand grasping of virtual objects. This standing problem raises the following question: can freehand grasping be performed in an interaction task without requiring users to be highly accurate in size and position estimation of virtual objects and still be usable? This chapter will answer this question through implementing user-based grasping tolerances to assist users in freehand reach to grasp interaction with virtual objects. These tolerances are essentially the user errors found in Study 1 (Chapter 5), and are applied within two configurations, namely absolute to the positioning of the object in x,y,z space and secondly relative to the z plane positioning of the object only. This study will only focus on user task completion times and usability of freehand grasping with the application of the two types of tolerances, thus the accuracy of grasping in measures of GAp and

$GDisp$ will not be part of the analysis. This is done for two reasons. Firstly, the user tolerances that will be applied are unique values for each position (i.e. absolute tolerances) or z plane (i.e. average tolerances) that differ in magnitude for each virtual object (cube and sphere), thus measuring and comparing grasping accuracy in between different tolerances would be invalid. Furthermore measuring grasping accuracy would potentially result in a false interpretation of the results as performance will naturally be dependent on, and dictated by, the size of the tolerances applied. Secondly, this study will aim to illustrate whether freehand grasping can be usable to complete different interaction tasks (that have a start and an end) without having to be highly accurate in virtual object size and position estimation to trigger and finish the interaction task, thus the accuracy of users in this study is not needed.

8.1.1 Interaction Tolerances

Few studies in current research have aimed to improve interaction with virtual objects using interaction tolerances, for example, Hough et al. (2015) presented adaptive bi-manual interactions with virtual objects in an AR environment. Their work aimed to improve the plausibility of interaction in an AR scene using interaction offsets, where the size and position of virtual objects are adapted according to user hand movements. Using interaction tolerances to improve freehand grasping of virtual objects in AR environments, and assessing the impact of grasp phases on user grasping accuracy remain largely unexplored.

This chapter will address the problem of inaccurate size estimation of virtual object size, and investigate how this problem can be negated using user-based tolerances that will assist users in performing freehand grasping without needing to be highly accurate in size and position estimation of virtual objects. The impact of user-based grasp tolerances and grasp phases on freehand grasping performance will be measured using task completion time and the usability of this proposed method will also be addressed using the standardised System Usability Scale (SUS). Section 8.2 [page 174] firstly defines the two types of grasping tolerances used in this study, and outlines the design of the two experiments in this study in terms of the conditions under test, participants recruited and the experimental protocol. Section 8.3 [page 181] then discusses the data collected in the two experiments of this study that compare the two conditions: absolute and average user based tolerances, and provides a comprehensive analysis of the interaction performance in terms of task completion time and usability results. Finally Section 8.4 [page 185] provides the conclusions drawn from this study and a summary of the key outputs.

8.2 Study Outline

8.2.1 Interaction Grasping Tasks

In the past 3 studies (Chapters 5-7) users were required to reach for and then grasp static virtual objects and decide when they felt confident that they have securely grasped the object. The grasping phase analysed in the past three studies (Chapters 5-7) was the pre-load phase, that follows the initial reaching phase for the object. The pre-load phase is by definition the stage at which the user forms a stable grasp around the object after correcting any errors in the grasp strategy (e.g. grasp type, posture and position) that may have occurred during the initial reaching phase (see Section 3.5.2 [page 43]). Following this phase is the transition phase where the user performs the aim of the grasping movement such as move the object, lift it or use it as a tool for manipulation. Due to the lack of haptic feedback in exocentric AR and the lack of virtual object weight in AR in general, this pre-load phase in the previous three studies ended once the users decided that they are satisfied with their grasp. In this study the user based tolerances from Chapter 5 are used to form this pre-load phase during a grasping movement to separate the pre-load phase from the following transition phase, where the pre-load phase ends once the user triggers the interaction by performing a grasp that is within the tolerances applied in terms of GAp and $GDisp$. Thus the tolerances applied in this study are only applied in the pre-load phase of a grasping movement. This will potentially aid in achieving the overall aim of this study, that is enabling users to form a natural grasp in an interaction task with virtual objects that are normally grasped in real environments.

Based on this, the tasks in this study will require users to grasp (within the tolerances applied) and then move an object from a starting location to a target location in a two step grasping movement. Similar to the three previous studies in this thesis (Chapters 5-7), this study will only analyse the results in the pre-load phase using task completion and usability, this is essentially the time it takes users to trigger the interaction (i.e. complete the pre-load phase by performing a grasp that is within the tolerances applied). Thus the term "task" in this study refers to pre-load phase of grasping only, and not the following transition phase. Even though users in this study also completed the consequent transition phase (in order to complete and end the interaction task), this phase is not included in the analysis for clarity in relation to previous studies, and as no statistically significant differences were found between tolerance or object types in the time it took to move virtual objects to a target location as illustrated in the published version of this study (Al-Kalbani et al., 2017). For this reason, task completion time that is used to assess usability and user performance in this study in the pre-load phase only is representative of the whole interaction as users spent the majority of the time in the reaching and pre-load phases of grasping movements.

This will evaluate natural freehand grasping in an interactive task and its usability in completing interaction tasks for interactive AR systems. Using grasping for these tasks and par-

ticularity moving virtual objects from one location to another can potentially offer a usable and natural interaction method to be used in wide range of applications such as engineering/manufacturing assembly tasks (e.g. Evans et al. (2017)), medical/surgical training (e.g. Frajhof et al. (2018)) or education applications (e.g. Khan et al. (2018)) using AR.

8.2.2 Interaction Offsets

Two interaction parameters, Grasp Aperture Offset (aO) and Grasp Displacement Offset (dO), are introduced to investigate their effect on freehand grasping interaction. Grasp offsets (aO and dO) in this study are based on user errors in Study 1 (Chapter 5), where users were instructed to match the size and position of static virtual objects in different locations in 3D space using a medium wrap grasp (see Figure 4.2a [page 50]). Thus the grasp offsets are essentially the user errors found in Study 1 (Chapter 5).

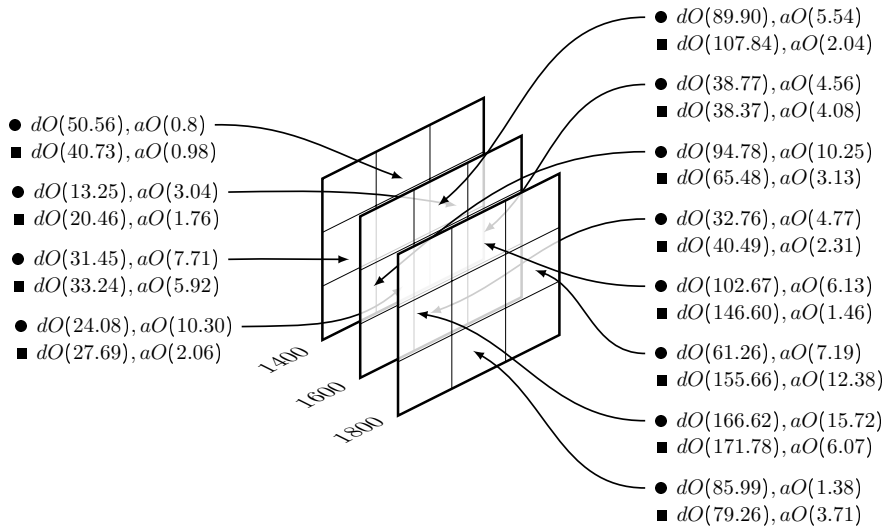
aO is defined as an interaction tolerance ± user Grasp Aperture (GAp) that defines a GAp range within which an interaction can occur. Likewise, dO is defined as an interaction tolerance ± user Grasp Displacement ($GDisp$) that defines a $GDisp$ range within which an interaction can occur. In this study, the grasp measurements (GAp and $GDisp$) are required to be within the range of both offsets (aO and dO) for the grasping interaction to be triggered, where larger values of aO and dO facilitate object grasping such that users need be less precise with grasp placement and aperture. Thus user errors found in Chapter 5 (Study 1), where users misjudged the position and size of the virtual objects presented, are recreated as offset parameters to test their impact on freehand grasping interaction, using absolute and average tolerances against task completion time and usability.

Absolute Tolerances

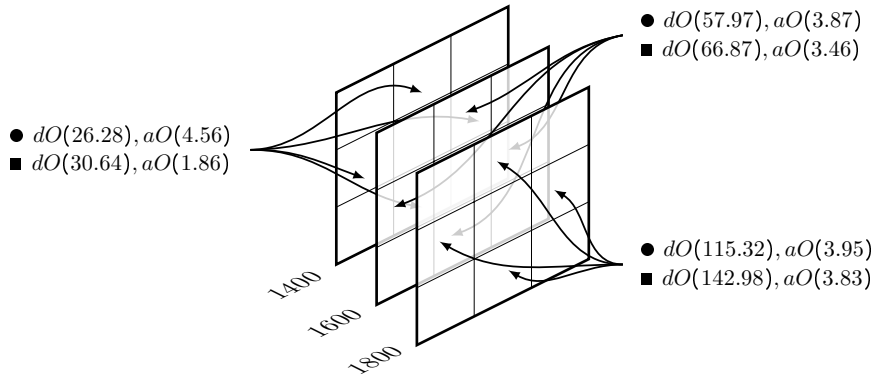
Absolute tolerances are aO and dO that are unique to a single object position. Absolute tolerance are different for each position and virtual object type (cube or sphere) in each position. These tolerances are small for being unique to a single position, and as such users will be required to be accurate in their grasp performance in order to be able to complete the pre-load phase of their grasping movement before moving the object. Evaluating the impact of absolute tolerances will potentially provide insights on how effective personalised or unique tolerances are for assisting freehand grasping. This can potentially be useful for applications where high accuracy in grasping is required.

Absolute tolerances are shown in Figure 8.1a [page 176] in four positions in each z plane (1400mm, 1600mm and 1800mm), thus 12 positions in total. These positions are the same positions in which users performed grasping in Study 1 (Chapter 5). The 12 positions represent the starting positions of virtual objects in each of the four tasks used in this study (see Figure 8.2 [page 177]). Given the need for an interaction task across a certain distance to have a starting and a target location, evaluating every position that was used in Study 1 (Chapter 5) is not feasible. For this reason, four tasks are evaluated that represent two different interaction directions (horizontal

and vertical) (see Figure 8.2 [page 177]).



(a) Absolute tolerances in mm in all the 12 positions of the tasks in this study across all z planes. ● represents spheres and ■ represents cubes



(b) Average tolerances in mm across all z planes. ● represents spheres and ■ represents cubes

Figure 8.1: The two types of user based tolerances assessed in this Study. 8.1a: Absolute tolerances that are position based (unique to each position), and 8.1b: that are area based (per each z plane)

Average Tolerances

Average tolerances are means of the individual aO and dO offsets in all nine positions in a single z plane. Average tolerances are shown in Figure 8.1b [page 176] in the four starting positions in this study for each z plane (1400mm, 1600mm and 1800mm). Average tolerances are different for each z plane and virtual object type (cube or sphere). In contrast to absolute tolerances, average tolerance are larger and more general, this will enable users to be less precise in their grasp size and positioning in order to trigger the interaction (i.e. complete the pre-load phase). Evaluating average tolerances will potentially illustrate whether having more generalised, and larger,

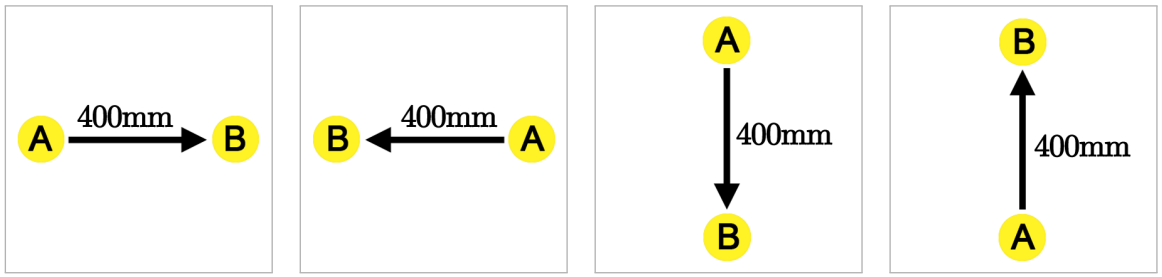
tolerances is more usable than unique absolute tolerances. This can particularly be useful for applications where speed in task completion is most important in an AR system.

8.2.3 Design

Two experiments were conducted in this study using the baseline setup outlined in Chapter 4 (see Figure 7.1a [page 140]):

- Experiment 1 to quantify the influence of absolute tolerances on grasp performance in terms of task completion time and usability given changes in object position in x,y and z space and object type
- Experiment 2 is a replication of experiment 1 to quantify the influence of average tolerances on on grasp performance in terms of task completion time and usability given changes in object position in x,y and z space and object type

A 2 × 2 × 3 × 4 repeated measures (within-subjects design) was used, with two primary conditions: absolute and average tolerances, four reach to grasp tasks (see Figure 8.2 [page 177]) and two objects (cube and sphere). Every permutation of tasks was randomly presented to participants to exclude potential learning effects. In total, each participant completed 2 (objects) × 2 (repetitions) × 3 (z planes) × 4 (tasks) = 48 trials and 1140 grasps in total (48 trials × 15 participants × 2 tolerance types). Findings from both experiments are compared to test the influence of user-based tolerances on grasp performance in terms of task completion time and usability, thus in this chapter results from both experiments are presented together and not independently for each experiment to analyse the impact of the primary condition in this study that is user based grasp tolerances.

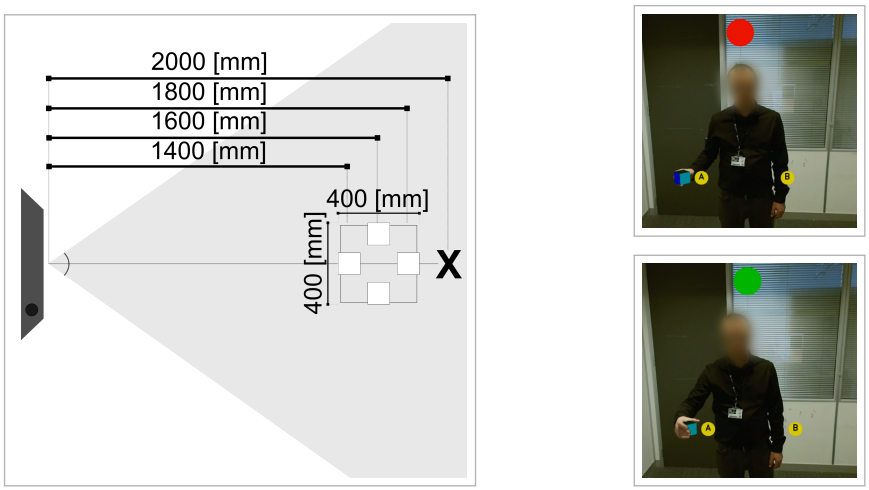


(a) Task 1 (b) Task 2 (c) Task 3 (d) Task 4

Figure 8.2: The 4 tasks that participants completed in this study. Letter A: starting location of the virtual objects. Letter B: target location. Arrows: motion direction. Distance between starting and target locations was constantly 400mm. Results in this study are only analysed in the reaching and pre-load phases of interaction (i.e. point A)

As shown in Figure 8.3a [page 178], the physical configuration in this study is identical to the one used in Chapter 5 (Study 1), where the main difference between the two setups is the pre-

sentation of virtual objects in 12 starting positions only that represent the four tasks evaluated in this study.



(a) Experiment setup

(b) User visual feedback

Figure 8.3: Exocentric AR system in this study. 8.3a: experiment setup, where X marks the standing position of participants. 8.3b: Feedback presented to participants during reach to grasp tasks. A and B yellow circles are the starting (A) and target (B) locations and the red and green circles show the state of the grasp. Top: participant locating the object (circle remains red). Bottom: participant successfully grasped the object (circle turns green)

Alongside occlusion handling, an additional visual feedback cue was also presented in the form of a circle that turns green if the virtual object has been successfully grasped or remains red otherwise. This circle was used to clearly distinguish between the different grasp phases in this study, where a red circle refers to the reaching phase of a grasp (i.e locating an object), and a green circle refers to the pre-load and transition phases of a reach to grasp movement. Positions of the starting and target locations were shown as A and B in yellow circles (see Figure 8.3b [page 178]). The distance between starting and target locations was constantly 400mm throughout the test.

Conditions of both experiments are shown in Table 8.1[page 179], where experiment 2 (average tolerances) is a replication of experiment 1 (absolute tolerances). User grasping performance using a medium wrap grasp is measured against task completion time to test the impact of the two user-based tolerances proposed on grasp performance in terms of task completion time and usability. To represent grasp performance independent of additional rendering, for both experiments, the baseline objects which have not undergone complex rendering and represent a simple abstract shape are used.

Hypotheses

$H_{4.1}$: Grasp tolerances (absolute and average) have no effect on task completion time and usability in grasping interactions

Table 8.1: Experiments 1 and 2 conditions, where x is measured from the centre of the sensor, y from ground and z from sensor

Experiment 1 and 2				
Condition	**Levels**			
Object Position (x, y) [mm]	Task 1	Task 2	Task 3	Task 4
	-400, 1250	400, 1250	0, 1650	0, 850
	* 4 starting positions were repeated in each z plane (1400mm - 1600mm - 1800mm), resulting in 12 positions in total			
Tolerances	Absolute (Experiment 1) and Average (Experiment 2)			
Object Type	Cube and Sphere			

8.2.4 Participants

30 right handed participants ranged in age from 21 to 64 (M = 33.97, SD = 9.84), in arm length from 480mm to 660mm (M = 557.07, SD = 40.64), in hand size from 130mm to 200mm (M = 185.23, SD = 14.26), in height from 1570mm to 1940mm (M = 1754.87, SD = 90.59) and 7 were female and 23 male. Taking into account balance in hand size, arm length, gender, age and height, participants were separated into two groups of 15 for the two experiments.

8.2.5 Statistical Analysis

Due to the format of the data collected being non-parametric and not normally distributed, statistical significance between the two independent groups in this study is tested using a non-parametric Mann Whitney-U test (Mann and Whitney, 1947) with an alpha of 1% comparing the two conditions (absolute and average tolerances).

8.2.6 Protocol

This study followed the baseline experiment protocol outlined in Section 4.2.4.2[page 62] prior to collection of data.

Participants underwent initial training of the medium wrap grasp on real and virtual objects, followed by training on two reach to grasp tasks. The test coordinator explained the procedure between each block of tests (i.e cube and sphere), and participants were allowed to rest before the presentation of every object. Each experiment was formed of a 5 minutes training/instruction session, 10 minutes of grasping a cuboid object, 5 minutes break and 10 minutes of grasping a spherical object (order of virtual objects counterbalanced).

After completing the test, participants were asked to fill in a usability questionnaire and a set of questions regarding their interaction to evaluate the ease of use and usability of our interaction system. The usability of the system was evaluated by a user satisfaction test using the System

Usability Scale (SUS) (Brooke et al., 1996). In order to further assess the usability of our proposed natural interaction system, participants were asked to complete an additional post-test questionnaire consisting of 6 close-ended questions regarding different aspects of their interaction (e.g. task difficulty, difficulty of each grasp phase, object type difficulty) (see Appendix C [page 222]). Questions/Statements were:

1. I found it easy to locate and successfully grasp objects
2. I found it easy to move objects to the target location
3. I have noticed that the virtual objects changed position in the x, y and z axes
4. Rate the difficulty of each task completed (Tasks 1, 2, 3 and 4)
5. Did you suffer from fatigue or pain during any of the tasks in this test?
6. Which of the two objects did you find easier to interact with?

8.2.7 Procedure

For the two experiments in this study, participants were instructed to locate the virtual object presented, successfully grasp it in the starting location (A), that corresponds to the positions illustrated in Figures 8.1a and 8.1b, using a right-handed medium wrap grasp and then move it in a controlled manner (i.e. straight line) to the target location (B) in the shortest time possible. This task-based design in this study covers three subsequent grasp phases: interaction starts with the reaching phase (i.e locating the object), this is followed by the pre-load phase that forms a stable grasp and corrects grasping errors in structure or force for grasping physical objects (i.e when the circle turns green in this study), and finally the transition phase where users move the object from a starting location (A) to a target location (B). Note that given the focus of this thesis on the reach to grasp movements, analysis of results in this chapter will only be focused on the first two phases of reach to grasp movements (i.e reaching and pre-load), and their impact on grasp performance in terms of task completion time and usability.

Before interaction, an object (cube or sphere) appeared on the feedback monitor in different positions depending on the task being presented (see Figure 8.2 [page 177]). The object sizes that had the lowest mean difference between GAp and object size found in previous studies in this thesis were chosen for this analysis (80mm for cubes and 70mm for spheres) and were unchanged throughout the two experiments in this study.

During the interaction, the time spent by participants in locating and moving virtual objects from an initial position to a target location was recorded. Tasks ended automatically once the target location was reached.

8.3 Results

8.3.1 Results - Completion Time

1400mm Z plane

A statistically significant difference in task completion time between absolute and average tolerances in different positions was found for cubes (U = 1.00 $\times 10^8$, p < 0.01) and spheres (U = 9.23 $\times 10^7$, p < 0.01).

1600mm Z plane

A statistically significant difference in task completion time between absolute and average tolerances in different positions was found for cubes (U = 1.01 $\times 10^8$, p < 0.01) and spheres (U = 1.25 $\times 10^7$, p < 0.01).

1800mm Z plane

A statistically significant difference in task completion time between absolute and average tolerances in different positions was found for cubes (U = 1.70 $\times 10^7$, p < 0.01) and spheres (U = 6.04 $\times 10^6$, p > 0.01).

8.3.2 Analysis - Completion Time

As shown in Figures 8.4a [page 183] and 8.4b [page 183]), in the reaching and pre-load phases for locating virtual objects before triggering the interaction users were faster using average tolerances, where a lower mean task completion time using average tolerances was found in 18 out of the 24 trials in this study (see Table 8.2 [page 182]). Task completion time in this study was largely dependent on the magnitude of tolerances applied to the grasping interaction, and this user preference for average tolerances in terms of completion time can be attributed to the difference in magnitude of tolerances between the two conditions (see Figures 8.1a [page 176] and 8.1b [page 176]).

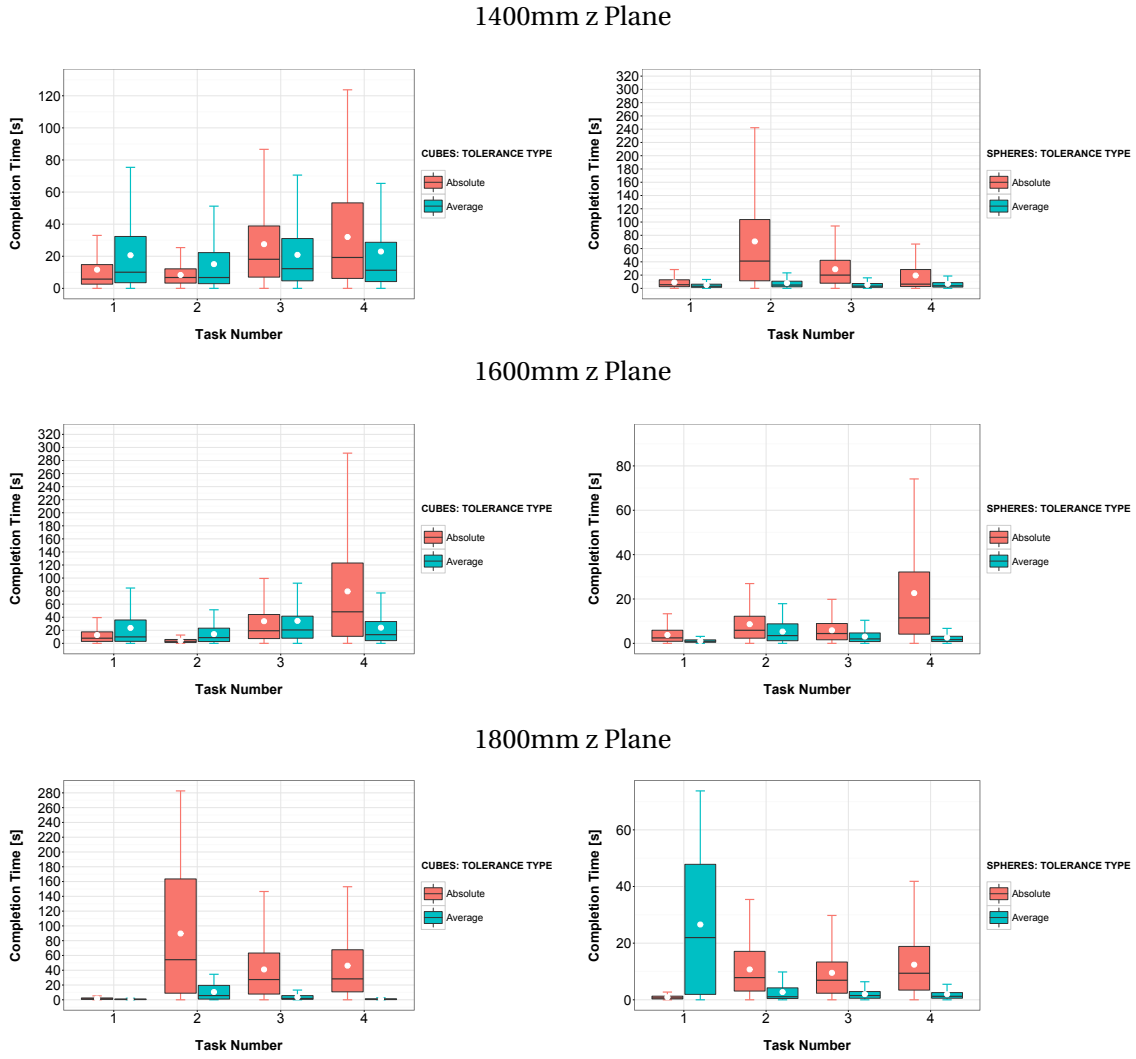
Participants spent the majority of the total interaction time in the locating phase of a grasp, where the time spent moving virtual objects to target locations was comparable between the two tolerance types (absolute and average). This shows that locating objects successfully in 3D space still remains the most challenging stage of freehand grasping in an exocentric AR environment.

8.3.3 Usability Analysis

A System Usability Scale (SUS) (Brooke et al., 1996) showed a mean score of 62.83 (SD = 18.68) for the absolute tolerances, while the SUS mean score for the average tolerances was 70.16 (SD = 14.04). The results can be labelled as “Ok” for absolute tolerances and “Good” for the average tolerances (Bangor et al., 2009*b*). These results may be linked to total completion time. Total completion time was higher for the absolute tolerances condition across the majority of tasks in all the three z planes in this study. User comments using post test questionnaires below provide

Table 8.2: Descriptive Statistics of Task Completion Time (Mean ± SD). Significant differences ($p < 0.01$) between absolute and average user based grasp tolerances are marked with (*)

1400mm Z Plane								
Object Type (Size)	Tasks							
	Task 1		Task 2		Task 3		Task 4	
	Absolute	Average	Absolute	Average	Absolute	Average	Absolute	Average
Cubes (80mm)	11.65 ± 13.78*	20.64 ± 22.66*	8.35 ± 6.40*	15.10 ± 17.64*	27.54 ± 27.97*	20.89 ± 21.14*	32.01 ± 31.31*	22.94 ± 28.29*
Spheres (70mm)	8.94 ± 8.92*	5.10 ± 5.32*	70.78 ± 77.60*	7.95 ± 8.30*	28.86 ± 27.51*	5.35 ± 5.38*	19.37 ± 25.08*	6.87 ± 7.54*
1600mm Z Plane								
Object Type	Tasks							
	Task 1		Task 2		Task 3		Task 4	
	Absolute	Average	Absolute	Average	Absolute	Average	Absolute	Average
Cubes (80mm)	12.73 ± 13.60*	23.67 ± 29.06*	4.10 ± 3.48*	14.18 ± 13.98*	33.93 ± 38.48	34.46 ± 38.87	79.60 ± 85.26*	24.02 ± 27.18*
Spheres (70mm)	3.79 ± 3.53*	1.11 ± 0.84*	8.64 ± 8.33*	5.30 ± 4.88*	5.89 ± 5.11*	3.05 ± 2.88*	22.65 ± 25.24*	2.59 ± 2.67*
1800mm Z Plane								
Object Type	Tasks							
	Task 1		Task 2		Task 3		Task 4	
	Absolute	Average	Absolute	Average	Absolute	Average	Absolute	Average
Cubes (80mm)	1.67 ± 1.34*	0.73 ± 0.60*	89.69 ± 89.24*	10.55 ± 10.94*	41.07 ± 41.41*	3.58 ± 3.92*	46.20 ± 46.37*	1.03 ± 1.06*
Spheres (70mm)	0.92 ± 0.83*	26.58 ± 24.22*	10.72 ± 9.12*	2.73 ± 3.29*	9.53 ± 9.12*	2.00 ± 1.84*	12.40 ± 10.83*	1.96 ± 2.02*



(a) Cubes

(b) Spheres

Figure 8.4: Task completion times for the reaching and pre-load phases of interaction in the four tasks in the three z planes in this study (1400mm, 1600mm and 1800mm). 8.4a: Cubes. 8.4b: Spheres. White points on boxplots indicate the mean completion time across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

general subjective insights regarding their experience in grasping virtual objects using dual view visual feedback, however these insights may not be directly representative of user performance and accuracy during interaction as these subjective responses were not measured against performance in this work.

Out of 15 participants in the absolute tolerances condition, four (26.66%) reported that the cube was easier to interact with, while 10 (66.66%) of the remaining reported it was the sphere. One participant remained undecided. In the average tolerances condition, three participants (20.00%) reported the cube as easier to interact with, while 10 (66.66%) reported it was the sphere and two remained undecided. One user also commented that "the sphere appeared easy compared to the cube". This preference for the sphere may be linked with measured task completion time and tolerance magnitudes. Tasks with the cube virtual object took longer to complete than the sphere in both conditions (absolute and average), with the exception of the 1400mm z plane (in absolute tolerances) and the 1800mm z plane (in average tolerances). aO for spheres were also larger than aO for cubes across all tasks, conditions and z planes, with the exception of task 3 in the 1800mm z plane and task 4 in the 1400mm z plane in the absolute tolerances condition (see Figures 8.1a [page 176] and 8.1b [page 176]).

For the usability analysis participants were asked to give answers to a set of specific close questions, a five-point Likert scale was used to record this additional usability feedback. Participants were asked to report on how easy they found to locate the object in the space, with 1 being extremely difficult and 5 extremely easy. Participants in the absolute tolerances condition scored this with a 2.93 out of 5. One user in the absolute tolerances condition commented that "the system was easy to learn, but grasping caused frustration", and another also stressed that it was "very frustrating when I could not successfully grasp the object". This shows that low tolerances can potentially cause frustration if it leads users to spending a long time to complete a grasping interaction task. On the other hand, the average tolerances condition scored a higher score of 3.47 out of 5. A user in the average tolerances condition also commented that "the tests were straightforward and easy to achieve".

When asked about how easy was to move the object, scores were comparable for both conditions: 4.86 (absolute tolerances condition) and 4.60 (average tolerances condition). This was expected as the tolerances in this study were only applied in the reaching and pre-load phases of a grasp and not in the transition phase. Users in both conditions (absolute and average) also commented that objects were "easy to move", and another user also commented that moving occurred "very easily and quickly". These results, as the ones introduced previously, may be linked to completion time, as it is one of the objective usability metrics linked to the subjective usability experience, the faster a participant can complete a task generally the better is the user experience (Albert and Tullis, 2013).

Participants were also asked to report if they felt any fatigue or frustration during the performance of the test, this was included in the questionnaire to better understand the impact of tolerance magnitude on user fatigue, as very small tolerances (e.g. absolute tolerances in this

study) can prolong task completion time as it requires users to trigger a very narrow tolerance using their grasp aperture and position. This can be particularly challenging given the lack of haptic feedback in exocentric AR. 9 participants in the absolute tolerances condition reported to have experienced this during the test while 5 reported the same in the average tolerances condition. Vertical movement tasks (3 and 4) were the ones linked to causing more discomfort. One user described vertical tasks as "considerably more difficult". These tasks are the ones with the longer completion times for the absolute tolerances group.
Users also provided additional comments regarding the overall design of the system and interaction, where one user suggested that "audio feedback for grasp status would be more desirable. Visual feedback (i.e. the additional circle) was less effective as my eyes were busy". Using audio feedback can be considered in future work, and can potentially offer an alternative solution as a feedback method for the status of a grasp, and can also lower the cognitive load for users by reducing the amount of visual information that users need to focus on during grasping movements. Another user also suggested that "the depth of the object would be good if there was some other visual indicator (such as light, drop shadows ect.". This further emphasises the importance of drop shadows that were used in Study 3 (Chapter 7) in improving depth estimation, and also in making depth changes more perceptible. Future work can look into using drop shadows alongside user based tolerances to assist users in freehand grasping.

Hypothesis - Revisited

$H_{4.1}$: Grasp tolerances (absolute and average) have no effect on task completion time and usability in grasping interactions: **Rejected** as statistically significant results were found showing that using absolute or average user based grasp tolerances in grasping virtual objects that change in position has a significant effect on task completion time.

8.4 Conclusions

This chapter presented a first study looking into the application of user-defined tolerances for improving natural freehand interaction between a user and a virtual object in an exocentric mixed reality environment. Based on previous user interaction analysis in Chapter 5 (Study 1) two definitions for user freehand tolerances are analysed, namely a Grasp Aperture Offsets (aO) and Grasp Displacement Offsets (dO). The offsets are applied within two configurations on different virtual objects (a cube and a sphere). The first configuration is based on the absolute positioning of the object in x,y,z space (absolute tolerances) and the second is based on average offsets for the object z plane in interaction space (average tolerances). User grasping performance in this study was measured using task completion time for grasping virtual objects that change position in the x, y and z axes.
This study showed that the application of average user tolerances was found to improve task completion time in the reaching and pre-load phases of interaction, thus offering an alternative

solution to drop shadows and dual view visual feedback in improving task completion time and usability in freehand grasping if drop shadows are not valid or feasible to use. This was reported across all interactions with the exception of two conditions where the absolute aO and dO were greater. The results further show that faster completion times were recorded for users interacting with spheres rather than cubes and again this is attributed to the increased aO and dO for the sphere object. This shows that for improving interaction usability and task completion then defining a more generalised average tolerances model based on object positioning can have an influence on task performance, thus increasing the aO and dO can lead to increased task performance. However, the impact this has on object interaction plausibility should be considered in future analysis. In addition, attention should be given to the speed/accuracy trade-off that is associated with the two tolerance types in this study (absolute and average) before implementing them.

Furthermore, the usability analysis found an increased usability when the average tolerances are applied (70.16 (SD = 14.04) compared to 62.83 (SD = 18.68) for absolute tolerances). This shows that the application of tolerances can assist users in grasping in task-based interaction without requiring users to be highly accurate in virtual object size and position estimation. In addition, users reported a preference when interacting with spheres rather than cubes, this could again be attributed to the increased aO and dO for the sphere object. Considering discomfort and frustration in the interaction, when questioned post-test, the users report on two conditions which caused frustration, namely tasks 3 and 4. These tasks relate to the longer task completion times and small tolerance values and did result in users adapting their grasp away from the defined medium wrap grasp to complete the interaction, thus showing similar deviation from the grasp type people were instructed to perform as found in previous studies in this thesis.

In Chapter 9 a discussion of the findings in Chapters 5-8 will be presented. Key findings in this work will be compared with current research in AR. Recommendations based on the findings from the four independent user studies in this thesis will also be drawn to aid in the development of more usable natural freehand grasping AR systems in the future. In addition, these recommendations will also present guidelines for implementing and improving freehand grasping accuracy in AR environments for interaction designers and the research community. Finally, the next chapter will also revisit the key parameters that impact physical grasping, and the transferability of these parameters to AR environments will be assessed based on the findings from the four user studies in this thesis.

Chapter 9

Discussion and Recommendations

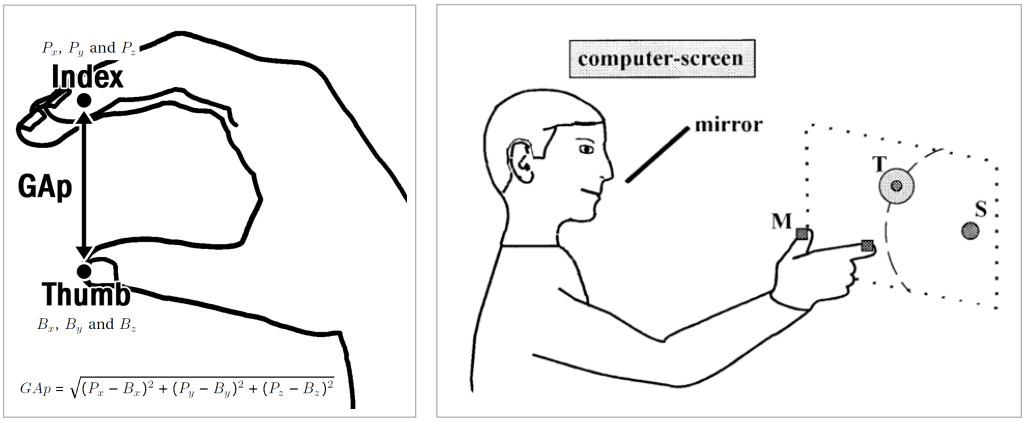
To complement the findings discussed across the four user studies in this thesis (Chapters 5 - 8), this chapter will present a discussion of transferable findings in this work and its implications on the research community and future AR applications in general, along with recommendations and routes for further research. Realising the accuracy and problems of freehand grasping can be used to aid in the development of future AR systems, and more importantly aid in bridging the gap between reality and virtuality by allowing users to use natural interaction techniques, similar to the ones used in real life, to interact with virtual objects. Evans et al. (2017) recently assessed the use of Microsoft's HoloLens in AR engineering assembly tasks, and pointed out that interaction is generally impeded by the lack of support of freehand interactions. This shows that further research is needed in this field to address the need of more natural interaction techniques rather than focusing solely on the capabilities of current available AR hardware, as natural interaction can potentially mitigate this problem. In addition, understanding the usability and accuracy of natural grasping in AR can provide interaction designers with various interaction solutions through utilising the unique interplay between the fingers and potentially access the 33 different physical grasp types, this unique interplay is not necessarily present in other gesture-based interaction techniques and could aid in increasing attachment and connection when interacting with virtual objects. This is particularly true in AR environments where real and virtual objects can coexist in the same environment, thus it seems plausible to use natural interaction techniques such as grasping in these environments where users can actually see their real hand and its interactions.

Sections 9.1 [page 188] and 9.2 [page 192] will discuss the main findings in this thesis based on the two grasping parameters used to analyse the accuracy and usability of freehand grasping, namely Grasp Aperture (GAp) and Grasp Displacement ($GDisp$). Finally Section 9.3 [page 199] will present usability and design recommendations for future AR systems that are based on the findings in the four user studies in this work.

9.1 Grasp Aperture (GAp)

GAp in this work measured the Euclidean distance between the user's index finger and thumb, and is a widely used metric to assess user performance in physical grasping (Edsinger and Kemp, 2007; MacKenzie and Iberall, 1994). Understanding the accuracy of GAp is important for measuring user performance in real world applications, especially when the hand opening is an indication for user ability. For example, Luo et al. (2006); Mei et al. (2017) used GAp in an AR system to determine user progress in post stroke rehabilitation through comparing user performance against desired outputs in grasping movements under the supervision of remote medical professionals. GAp in this work aimed to quantify user grasp accuracy in terms of virtual object size estimation in a natural user interface. Furthermore GAp also provided valuable insights regarding user behaviour and preferences in freehand AR grasping.

Users in this work constantly showed a working GAp range from 60mm to 80mm, even though they were presented with virtual objects that ranged in size from 40mm to 100mm (studies 1 and 2). This showed that GAp is not directly proportional to object size as is the case for grasping real objects where the size of the object is a key parameter in dictating the hand opening. Inaccuracy in virtual object size estimation using GAp is in alignment with findings in previous research. For example underestimation in virtual object size using GAp was found in the early work of Bock and Jüngling (1999) that measured grasp accuracy using finger trackers (on thumb and index finger) against grasp aperture changes (see Figure 9.1 [page 188]).



(a) Grasp Aperture (GAp) in this thesis

(b) Grasp Aperture (GAp) in previous research. Image courtesy of Bock and Jüngling (1999)

Figure 9.1: Measuring user grasping accuracy using GAp in 9.1a: this work, and 9.1b: previous research

Inaccuracy in the estimation of virtual object size in this work is attributed to the feedback method used being single view visual feedback, and the missing sensory information in the hand with regards to the virtual object in exocentric AR environments. This type of feedback did not allow users to visualise their full hand during grasping, thus users were not fully aware

of their grasp errors in terms of size estimation using GAp using only single view visual feedback. Even with the introduction of additional feedback cues such as dual view visual feedback in Study 2 and drop shadows in Study 3, users still showed a working GAp between 60mm and 80mm (see Figure 9.1 [page 189]).

(a) Single view visual feedback (b) Dual view visual feedback (c) Drop shadows

Figure 9.2: Visual feedback methods used in this work 9.2a: single view visual feedback using a single monitor, 9.2b: dual view visual feedback using an additional side camera and 9.2c: drop shadows alongside occlusion

This was surprising in the case of dual view visual feedback as users were able to visualise their full hand and aperture using the additional view provided. However, users generally focused more on correcting their grasp placement which was a more visible problem using the second view and not their hand opening. These findings show that inaccurate virtual size estimation using GAp can still occur even when using additional feedback modalities or methods. This has been illustrated in previous research that assessed GAp accuracy against virtual object size using multimodal feedback methods. For example Magdalon et al. (2011) assessed the impact of visual and haptic feedback on the kinematics of reach to grasp movements in virtual and real environments, where they used a 3D tracking system that placed infra-red emitting diodes (IREDs) on the head, trunk, arm, forearm and hand, and additional multiple trackers on the index, thumb, wrist, elbow and shoulder. Even though additional feedback modalities were used in their work to recreate the sensory information experienced during physical grasping, their work still reported on slower hand motion grasping of virtual objects with longer deceleration times, and also found overestimation of virtual object sizes. Furthermore, Bozzacchi and Domini (2015) also found a trend of decreasing GAp with increasing distance when grasping virtual objects using visual and haptic feedback modalities.

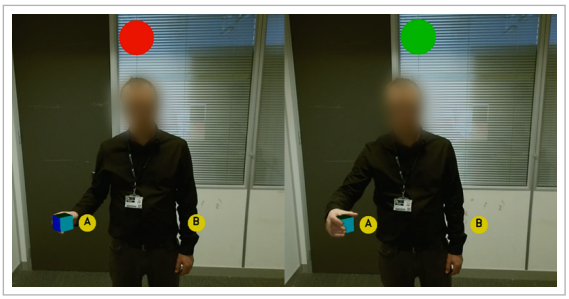
Two object types (cube and sphere) were used to assess grasp accuracy in terms of size estimation in this work. In studies 1 and 2 users were presented with different cube and sphere sizes (40mm, 50mm, 60mm, 70mm, 80mm and 100mm) to assess user accuracy in size estimation of virtual objects using GAp. Significant differences were found in grasping accuracy between cubes and spheres in this work, however users still performed GAp within the 60mm to 80mm

range for both objects. Users were generally more accurate in grasping cubes than spheres in this work. This can potentially be attributed to the shape of the cube that clearly showed users the different faces of the cube, that lead users to potentially perform grasps that complied with the geometrical features of the cube thus leading to more accuracy. Interestingly however, users indicated that grasping spheres was easier and more straightforward even though they were more accurate in grasping cubes. Users justified this preference for grasping spheres by indicating that spheres looked more like physical objects that one would normally find and interact with in real environments. One user commented that "the sphere object looked like a ball that can easily be grasped", as one would normally do in a real environment with a physical spherical object. Another user also said that "the sphere appeared easy compared to the cube". Users here were potentially using their personal experience and memory in their grasping strategy (Anderson, 1985), where the sphere for them felt like a real object such as a tennis ball. Using this previous experience in grasping spherical physical objects potentially made grasping the virtual sphere object seem perceptually easier to grasp, owing to the fact that previous experiences and memories can play a big role in perceiving different objects during grasping movements (see Section 3.3.2 [page 32]). Previous research by Swan et al. (2017) has also shown that users are more likely to interact more accurately with familiar objects than unfamiliar ones, as familiar objects can enable users to use the familiar size and shape of virtual objects as a cue for accurate size and distance estimation. It can be argued that cubes can also be labelled as "familiar", as one also has experience in grasping cubic physical objects in real environments such as match boxes. However for virtual cubes in this work, unfamiliarity refers to users potentially not having as much experience in grasping perfectly symmetric cubes (equal height, width and depth) in real environments such as the ones presented in this analysis. Further tests are required to assess the impact of virtual object shape on perception during freehand grasping to further understand this perceptual insight found in this work, especially as the true meaning of the term "familiar" that is used to describe objects can significantly vary from one user to another based on their previous experiences in grasping. For example future work can assess natural grasping of more complex objects that are familiar to users in real environments, such as objects with handles (e.g. mugs) or curved objects (e.g. bananas).

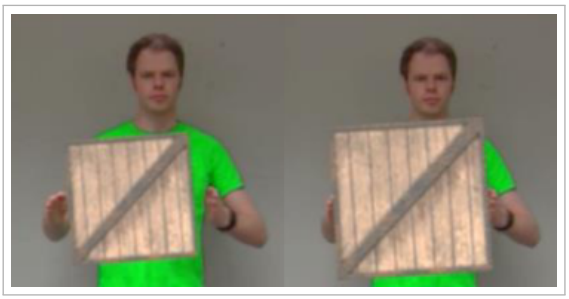
Grasp accuracy and stability in terms of object size estimation using GAp that is inherited to the pre-load phase for grasping real objects are not necessarily warranted for freehand grasping in an AR environment. The pre-load phase in grasping real objects always provides a stable grasp through overcoming any external forces and task constraints. However, findings across all the user studies in this thesis showed that users are generally inaccurate in size estimation in this pre-load phase. This shows that transferring the assumptions associated with grasping real objects to freehand grasping in AR environments is not suitable. This work alongside previous research that assessed grasping accuracy using GAp show that interaction designers should be aware of the potential discrepancies in virtual size estimation during grasping. These discrepancies using GAp present a fundamental problem for grasping interaction, and can potentially

be attributed to the size distortion that is caused by virtual object perception, and the inaccurate scaling of retinal information with increasing distance in AR environments.

The standing problem of inaccurate virtual object size estimation using GAp in this work raised two questions: is having the identical accuracy to grasping real objects necessary in AR environments especially with important elements such as object weight and tactile feedback missing? and is it possible to have usable grasping interaction without being highly accurate in terms of grasp aperture? Study 4 addressed these problems through aiding grasping interaction with two types of user-based tolerances in interaction tasks; average tolerances that are area based (i.e. per z plane) and absolute tolerances that are position based (i.e. x, y and z coordinates). The use of tolerances was inspired by previous research that presented adaptive bi-manual interactions with virtual objects (Hough et al., 2015), where the use of interaction offsets reduced interaction errors and improved the plausibility of adaptive bi-manual interactions with virtual objects in an AR environment. Their work firstly assessed user interaction errors using two metrics; Mean Distance to Object Surface and The Variability in Distance Between Hands. Quantifying user errors in their work using these two metrics then allowed improving user interaction accuracy through adapting the size and position of virtual objects according to user hand movements during interaction, this accordingly improved the overall interaction user accuracy and plausibility for third person viewers (see Figure 9.3b [page 191]).



(a) User based tolerances in this work to assist in freehand grasping



(b) Adaptive bi-manual AR freehand interaction based on user hand position and variability. Image courtesy of Hough et al. (2015)

Figure 9.3: Adaptive methods for assisting in freehand interaction with virtual objects in AR in 9.3a: this work, and 9.3b: previous research

Findings in this final study showed that the use of the more general average tolerances can negate this fundamental problem of having to accurately grasp virtual objects in terms of size estimation while maintaining an acceptable overall usability of the system. Application of user-based tolerances significantly reduced task completion time and made the overall interaction for users easier, while still requiring them to perform an actual grasp to trigger the interaction (see Figure 9.3a [page 191]). Application of average tolerances scored 70.16 (SD = 14.04) using SUS, and were found to be “Good” in terms of usability according to the usability rankings of Bangor et al. (2009*b*). Users also commented that freehand grasping interaction tasks with

the application of tolerances "were straightforward and easy to achieve". This shows that tolerances can aid in developing usable grasping interactions that allow users to use a natural interaction technique similar to the one used in real environments. Study 4 also highlighted the need to take into account the speed-accuracy trade-off associated with tolerances, this was evident by the significantly higher task completion times found under the absolute tolerances condition in this study. Thus if tolerances are very small to achieve higher grasping accuracy, it can significantly hinder usability. Future work should also investigate how the application of tolerances that aid in grasping interaction impacts the plausibility of interaction. Furthermore users also suggested that "the depth of the object would be good if there was some other visual indicator (such as light, drop shadows ect)", this can again be a route for further analysis where the impact of multiple methods for assisting freehand grasping (e.g. drop shadows and user based tolerances) on the accuracy and usability of freehand grasping in AR can be measured. Inaccurate virtual size estimation is a fundamental problem in AR environments as evident by findings in this work and previous research, and improving the type of user feedback regarding their GAp during interaction can potentially improve virtual size estimation. Users in this work suggested that using audio feedback to update the grasp status may be an alternative solution to additional visual cues (i.e. the coloured feedback circle used in Study 4 of this work) and is an approach that is also currently being investigated in current research for AR environments (Kimura and Sato, 2018). One user also suggested that "continuous feedback for grasp movements" can also be a route for further research, and can also be informative in giving users information regarding their hand structure, thus if users perform a grasp type, visual feedback can provide this information (e.g. using text). This can potentially make users more connected with their grasping performance.

9.2 Grasp Displacement ($GDisp$)

$GDisp$ was also introduced to be used alongside GAp, to provide information regarding the position and placement of a grasp in 3D space in relation to a virtual object. Using the placement of the Grasp Middle Point (gmp), Grasp Displacement ($GDisp$) is then calculated by measuring the position of the middle point of a virtual object (omp) from the gmp. This results in the distance from the middle point of the grasp to the middle point of the virtual object in the x ($GDisp_{\text{x}}$), y ($GDisp_{\text{y}}$) and z ($GDisp_{\text{z}}$) axes (see Figure 4.4a [page 52]). As shown in Figure 9.4 [page 193], using distance from a target location is a widely used metric in assessing user performance in AR environments (e.g. Swan et al. (2015); Chen and Saunders (2016); Kim and Park (2015)).

In this work $GDisp$ provided information regarding the position of a grasp in the x, y and z axes ($GDisp_{\text{x}}$, $GDisp_{\text{y}}$ and $GDisp_{\text{z}}$).

The first baseline study in this work (Chapter 5) showed that users significantly underestimated object position in the z axis. This underestimation highlighted the inherited problem of inaccu-

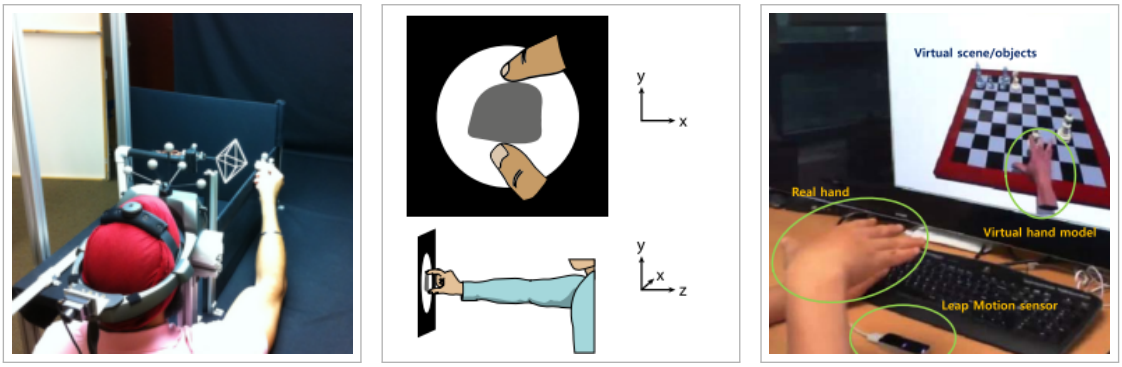
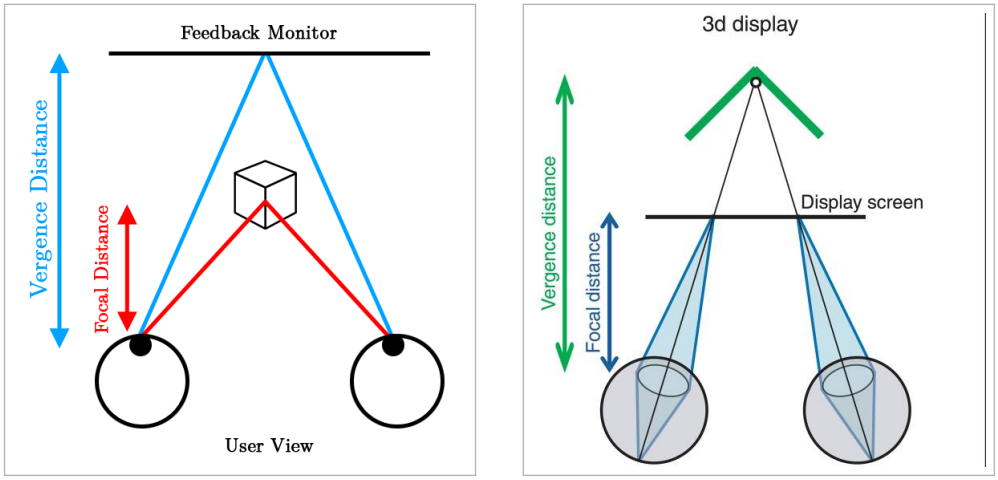


Figure 9.4: Distance from target metrics used to assess user accuracy in interaction and grasping in previous research. From left to right: images courtesy of: Swan et al. (2015), Chen and Saunders (2016) and Kim and Park (2015)

rate depth estimation in AR applications (Swan et al., 2017), that is largely influenced by the distorted perception of virtual objects in AR environments. In exocentric systems such as CAVEs, that are comparable to the setting used in this work, Bruder et al. (2015) found that user distance from the feedback screen directly impacts depth estimation accuracy. Inaccurate depth estimation is a fundamental challenge in AR for interaction designers, as blending reality and virtuality in an AR interface is a perceptual task where the interaction designer attempts to convince the human perceptual system that the virtual information presented is as realistic as the surrounding world (Billinghurst et al., 2015). In natural human vision, perception of size and distance information is inferred using various depth cues such as pictorial (e.g. occlusion), kinetic (e.g. motion perspective and parallax), physiological (e.g. vergence and accommodation) and binocular disparity (e.g. combining two views of the scene) cues (Drascic and Milgram, 1996). Accurate depth perception is possible in the real world owing to the fact that depth cues are almost always in alignment, in AR however perception of depth cues can be distorted due to missing or uncontrolled depth cues, and it is almost impossible to control all possible perceptual cues which can distort perception and directly affect task performance (Billinghurst et al., 2015). It is widely known that inaccuracy in depth estimation is caused by the accommodation - vergence conflict in wearable based AR systems (Swan et al., 2015). For example, users can experience an accommodation - vergence conflict when using stereoscopic displays (see Figure 9.5b [page 194]). This problem occurs due to a conflict in two physiological depth cues, namely vergence (rotations of eyes in opposite directions to focus on a specific depth) and accommodation (change of focal length of the eye where muscles attached to the lens of the eye contract and relax to perceive close and far objects), and occurs when the eyes converge on the virtual object that is viewed in two spatially offset views provided by the right and left eyes, but accommodate at a different depth that is usually the constant depth of the display (Kruijff et al., 2010). This problem is often associated with eye strain while using stereoscopic displays, however it is argued that the human eye is capable of adapting to this conflict (Kruijff et al., 2010) with multiple studies (Wade and Swanston, 2013; Kersten and Legge, 1983; Drascic and Grodski,

1993) showing that users did not suffer from painful side effects as is usually claimed. However, even if this conflict is not a major contributor to eye strain, it can still significantly impact perception in AR (Mon-Williams and Tresilian, 2000).



(a) Potential vergence - accommodation conflict in this work

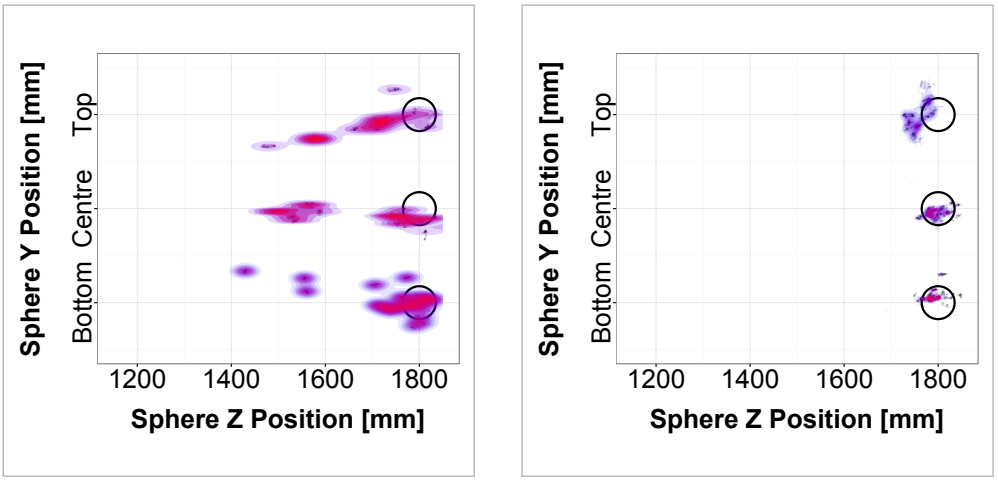
(b) Vergence - accommodation conflict in wearable AR systems. Image courtesy of Hoffman et al. (2008)

Figure 9.5: Perceptual vergence - accommodation mismatches in 9.5a: this work, and 9.5b: previous research in AR

In systems such as the one analysed in this thesis, where the environment (or user hand) is viewed directly (e.g. OST HMD or monitor based), an accommodation mismatch between real and virtual objects almost always occurs (Drascic and Milgram, 1996). This mismatch occurs as the accommodation distance to the virtual object is the distance between the eyes and the display, whereas the accommodation distance to the real object (e.g. user hand) is the distance between the eyes and the real object (the hand in this case). Thus for example if a user aligns their hand with a virtual object, this mismatch in accommodation distances serves as a strong depth cue that informs the user if the virtual object is in this specific position or not. This can potentially be the cause of inaccurate size and depth estimation found in this work in the natural exocentric setting used. One user commented that "I was sometimes looking at my hand in mid-air and not the feedback screen to see my grasp", this potentially caused a perceptual conflict that hindered size and depth estimation as looking at the hand would cause a vergence - accommodation conflict even if the object cannot be visualised in that particular position, the real object here is the real hand (see Figure 9.5a [page 194]). This problem can significantly impact task performance depending on the difference in accommodation distances (larger accommodation distances provide stronger depth cues and vice versa). There is currently a lack of research that is focused on understanding the impact of vergence and accommodation on user performance in natural user interfaces, and further analysis is required to fully understand the impact of this perceptual problem in exocentric AR systems. Improving depth estimation

in AR can make it a useful tool for applications such as image-guided surgery, manufacturing and maintenance that require high accuracy in judging the depth of virtual objects (Swan et al., 2015).

For freehand grasping in this work depth underestimation (i.e. user hand is closer to the sensor than the virtual object in the z axis) is attributed to the single view visual feedback used in an exocentric setup, where users were unable to visualise their interaction in the z axis. This problem of depth underestimation in the z axis was significantly improved using the dual view visual feedback that was assessed in Study 2, through providing users with a secondary view that made users aware of their grasp placement errors and accordingly enabled them to correct their depth estimation during the interaction. Figure 9.6 [page 195] shows an example of the significant improvement caused by the use of dual view visual feedback in one grasping task. Adding a second view for visual feedback was found to shift user gmp placement along the z axis closer to the 0 origin for all tasks, and significantly reduced the range of $GDisp_z$ for grasping both cubes (ranged from -41.73mm ± 29.17 to 34.88mm ± 15.62) and spheres (ranged from -40.36mm ± 18.98 to 47.52mm ± 47.09), in comparison to the significantly wider ranges that were found for single view visual feedback (from -220.12mm ± 181.69 to 5.94 ± 34.00 for cubes and from -233.25mm ± 174.06 to 20.55 ± 43.38 for spheres).



(a) Inaccurate depth estimation using single view visual feedback in Study 1 (Chapter 5)

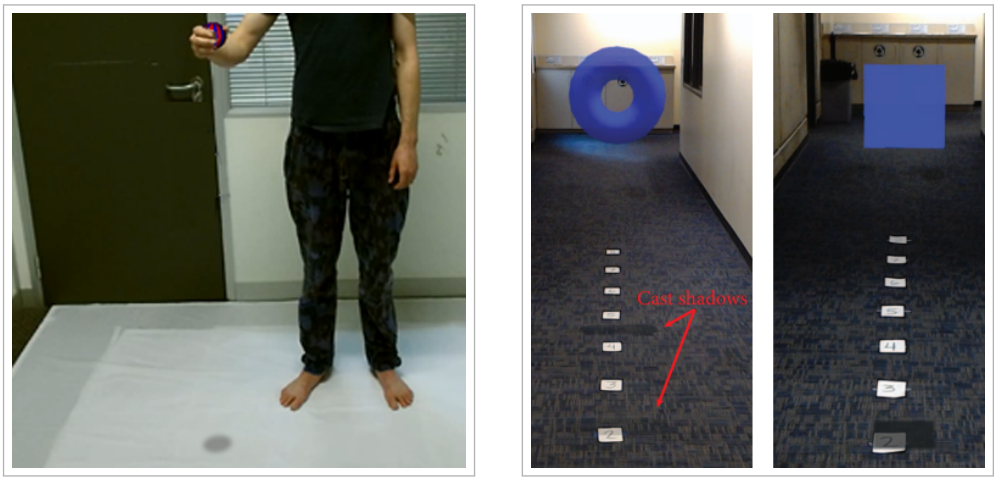
(b) Significantly improved depth estimation using dual view visual feedback in Study 2 (Chapter 6)

Figure 9.6: Example of improved depth estimation using dual view visual feedback

Early work of Hoang et al. (2011) is one of the few assessments in current research of the impact of multiple views on interaction with virtual objects in AR. Their work measured the impact of using different views from different cameras such as remotely located, head mounted zoom lens and tripod mounted zoom lens cameras on virtual object manipulation tasks. Their work illustrated that using multiple views in AR offers significant benefits such as zooming on regions of interest in an AR scene, additional viewing angels of virtual objects and higher precision in inter-

action due to the independent views of interaction that multiple views provide. Dual view visual feedback in this work scored 64.50 (SD = 13.43) using SUS, and was rated as "OK and marginally acceptable" according to the SUS rankings. In the post test questionnaire users labelled the additional side view as "necessary for accurate grasping", and that "it provides important information regarding the depth of the object". This shows that dual view visual feedback offers an effective solution for inaccurate depth judgements in exocentric AR environments, and can also be used in applications where high grasping accuracy is required. Dual view visual feedback also showed that there is a speed-accuracy trade-off associated that needs to be taken into account by interaction designers, as dual view feedback significantly increases task completion time and potentially the overall cognitive load during the interaction.

Drop shadows used in Study 3 also significantly improved depth estimation when used as an additional visual cue to aid in grasping. This is in alignment with current research in AR that illustrate that drop shadows have the largest impact on depth estimation (Diaz et al., 2017), in comparison to different rendering effects such as shading, cast shadows, aerial perspective and texture during perceptual depth matching tasks using the HoloLens (see Figure 9.7 [page 196]).



(a) Drop shadows in this work for natural freehand grasping

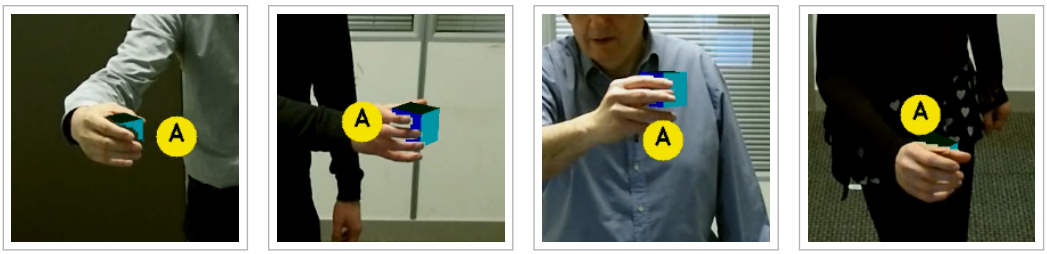
(b) Drop shadows assessed in previous research in AR. Image courtesy of Diaz et al. (2017)

Figure 9.7: Drop shadows assessment in 9.7a: this work, and 9.7b: previous research in AR

Drop shadows in this work have also improved the overall usability through significantly reducing task completion times. Drop shadows significantly reduced the range and variation of user task completion times, where it ranged from 6.86s $\pm$ 3.08 to 13.05s $\pm$ 9.16 for cubes and from 6.78s $\pm$ 1.22 to 10.26s $\pm$ 4.01 for spheres. These task completion times are significantly lower than the ones found using dual view visual feedback (8.47s $\pm$ 3.08 to 20.00s $\pm$ 19.83 for cubes, and from 7.53s $\pm$ 2.92 to 18.27s $\pm$ 11.86 for spheres). This significant impact of drop shadows in reducing task completion times was reflected in the comments of users in the post test questionnaire, where one user commented that "shadows help me to grasp the object fast and efficient".

Users generally stressed that drop shadows aided them in locating virtual objects even before interaction begins, and also stressed that drop shadows make position changes of virtual objects more perceptible. One user commented that shadows provided "a nice experience" that is "very promising". This is particularly useful in exocentric AR applications where users only have one view of their interaction. Shadows connects virtual objects to the real environment in AR applications resulting in more accurate depth judgements (Swan et al., 2017), and findings in this work further emphasise the importance of drop shadows to assist users in interaction in AR, especially when accurate depth judgements are required such as manufacturing and maintenance applications (Swan et al., 2015). Interestingly drop shadows were also rated higher by users in terms of usability when compared to the dual view visual feedback method, even though dual view visual feedback mitigated the problem of inaccurate depth estimation more significantly than drop shadows. SUS average score for drop shadows was 81.16 (SD = 11.56) that is rated as "GOOD and highly acceptable" (Bangor et al., 2009*a*), whereas for dual view visual feedback the SUS score was 64.50 (SD = 13.43) (rated as "OK and marginally acceptable"). This is potentially useful for future AR applications and interaction designers to be aware of, namely that additional feedback such as dual view visual feedback can significantly improve natural interaction performance, this can also lead to lower perceived usability. Thus choosing the best method for freehand grasping is largely dependent on the task and application requirements.

Similar to GAp, user based tolerances used in Study 4 were also aimed at enabling users to perform natural grasping without having to be highly accurate in grasp placement in relation to the position of the virtual object in the x, y and z axes. This method significantly negates the need for accurate depth estimation, and can potentially be useful for AR applications where accurate depth estimation is not required (e.g. entertainment AR applications). Findings in this study illustrated that freehand grasping with user based tolerances can be usable in interaction tasks without requiring users to be highly accurate in grasp placement in relation to a virtual object. Users also indicated in post test questionnaires that the application of tolerances during freehand grasping of virtual objects was "easy to learn", this is a positive indication that tolerances during grasping interaction can potentially aid novice users in feeling more connected to AR technology. Furthermore interaction tolerances also illustrated that users can interact with virtual objects in the natural form and posture that they would normally use in a real environment (i.e. grasping), regardless of the accuracy of their interaction (see Figure 9.8 [page 198]).
User based tolerances used in this work were effective in enforcing users to perform an actual grasp to interact with virtual objects, and to clearly distinguish between the the pre-load and transition phases of a grasping movement. This aided in accomplishing one of the main aims of this work; enable users to employ a usable and natural grasping technique that is naturally used in real environments to interact with virtual objects. Further work is required to assess how current grasp phases for grasping real objects can be transferable in a suitable manner to AR environments. This can potentially ease the process of assessing grasp accuracy and usability in each of these individual phases by aiding interaction designers in distinguishing between



(a) Task 1 (horizontal motion) (b) Task 2 (horizontal motion) (c) Task 3 (vertical motion) (d) Task 4(vertical motion)

Figure 9.8: User based tolerances aided users in performing natural grasping without requiring them to be highly accurate in estimating the positions and sizes of virtual objects

these phases when natural grasping is used, and also aid future systems in developing grasping interaction systems that take into account the impact of these separate phases on quality and accuracy of interaction.

Measuring $GDisp$ against the position of virtual objects in the x, y and z axes to quantify grasping accuracy, and taking into account the physical measurements of users in this work provided various valuable insights regarding user behaviour and preferences during natural interaction with virtual objects. For example in studies 1 and 3 where single view visual feedback was used, users showed higher accuracy in locating virtual objects in the z axis in the furthest z plane away from their bodies. This particular plane was at the extremity of the mean arm reach of users in these two studies, thus users did not have much room for error in this particular plane and were accordingly more accurate relative to the other two z planes. Users in this work were generally most accurate in grasping in the middle z plane that did not require extreme arm flexion (closest z plane to the body) or extension (furthest z plane from the body). This was evident for the dual view visual feedback method used in Study 2, where users preferred grasping in the middle z plane that was less physically demanding in terms of arm movement. Previous work of Chen et al. (2014) assessed user accuracy in reaching for physical and virtual objects in an exocentric CAVE environment, and found that users were inaccurate in estimating positions of far virtual objects away from their bodies due to awkward reaching postures (see Figure 9.9 [page 199]). Their work also emphasised the importance of considering user reach for virtual objects within a specific distance, and argued that considering the biomechanical features of users during reaching tasks can potentially allow for more natural interaction with virtual objects. This shows that arm movement is particularly important to consider when developing grasping interaction tasks that can potentially take a long time to complete (e.g. systems that are focused on accuracy and not speed).

Users in this work also showed a preference for object positions that were placed in the right and centre positions, this was mainly due to the fact that all users in this work were right-handed. A user commenting on left positions stated that “left positions required large shifts in positions” and hence were unnatural grasping movements. Previous research in AR (Piumsomboon et al.,

(a) Analysis of arm reach on freehand grasping accuracy in this work

(b) Impact of arm reach and posture on interaction accuracy in a CAVE environment. Image courtesy of Chen et al. (2014)

Figure 9.9: Arm reach assessment on interaction accuracy in 9.9a: this work, and 9.9b: previous research

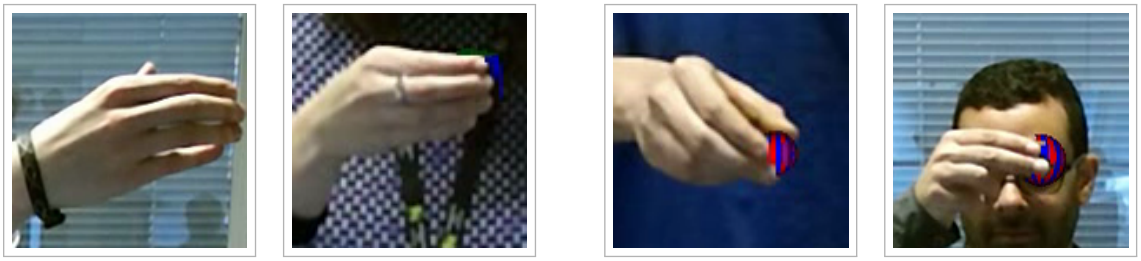
2013) has shown that users generally prefer using their dominant hand in fine motor interactions with virtual objects. Findings in this work are in alignment with previous research and show that placing virtual objects in the side of the dominant hand of users or not far from the side of the dominant hand (i.e. centre position) is more usable and user-friendly. This work also provided insights regarding the impact of the interaction direction (i.e. vertical or horizontal) on freehand grasping usability. In Study 4 where usability and task completion times were measured against 4 different grasping tasks, users indicated that vertical movement tasks were harder to complete. Users commented that "low and top positions were hard to get to", these low and top positions were essentially the two vertical tasks in this particular study (i.e. Tasks 3 and 4). This is in alignment with findings of previous research that assessed the impact of interaction direction on bi-manual interaction accuracy (Hough et al., 2015) and is again attributed to the level of arm movement associated with vertical tasks. Previous research in AR by Piumsomboon et al. (2013) also showed similar results when assessing gesture based AR interactions with virtual objects. Users in their work indicated that they found it harder to interact with virtual objects when they were required to lift their hands high during interaction. This again emphasises the importance of considering object position in relation to the arm and hand reach of users in AR applications.

9.3 Usability and Design Recommendations

Based on findings and user interaction trends found in this work, the following sections will provide recommendations for different aspects of natural grasping in AR environments.

9.3.1 Virtual Object Size

Findings in this work showed that users were most accurate in size estimation using GAp when grasping the 80mm cube and 70mm sphere, and these two sizes were then used for the remaining studies. However for smaller object sizes (i.e. less than 60mm in size), users were at times able to fully occlude the virtual object using a grasp. This can be problematic for freehand grasping as it can potentially hinder usability and accuracy (see Figure 9.10 [page 200] for examples).



(a) Fully occluding small virtual objects

(b) Changing grasp type or posture to show parts of the virtual object

Figure 9.10: Presenting users with small virtual objects can hinder usability, and can lead to users to 9.10a: fully occlude virtual objects or 9.10b: change their grasp type and posture to show parts of the virtual object during freehand grasping

Previous research in AR by Piumsomboon et al. (2013) also presented the same recommendation that users should not be able to fully occlude virtual objects, as this can hinder user experience when interacting in AR environments. One user in this work commented that he/she "was not sure how much of the object my hand should cover", this uncertainty during freehand grasping can be problematic in terms of accuracy and usability. Users in this work changed their grasp type and structure if the grasp they have performed fully occluded the virtual object presented, this was generally done by users to be able to visualise their interaction and grasp in relation to the object (see Figure 9.10 [page 200]). Thus users were more confident in their grasping interaction if they could see their grasp and, even if partially, the virtual object they are grasping. For example, if users are presented with a sphere that is 40mm in size, a medium wrap grasp could fully occlude the sphere, users may then change the medium wrap to a precision grasp using two or three fingers in order to show parts of the virtual object (see Figure 9.10 [page 200]). Based on the findings in this work, it is recommended that virtual object sizes should range from 60mm to 80mm in size for freehand grasping, the same working range performed by users in size estimation using GAp (see Section 9.1 [page 188]).

9.3.2 Grasp Type

Changing the type of a power grasp during interaction with real objects can lead to dropping or slipping of objects, thus a grasp type has to be constant in grasping real objects using a power grasp for it to be stable throughout all phases of a grasping movement (Napier, 1956). However,

this is not the case in freehand grasping as illustrated by the findings in the four user studies in this work, this is mainly due to the fact that sensory information (such as haptic feedback) and physical phenomena (such as gravity and mass) are not present in natural user interfaces such as the one assessed in this work. There are generally no consequences or physical constraints on users to keep their grasp type constant when grasping virtual objects. This raises the following question: should users be constrained by one grasp type when interacting with virtual objects in AR? Findings in this work show that users will most likely choose a grasp type that makes them believe that they are performing an accurate grasp for the conditions, tasks and feedback method used. Even if they are instructed to use a different grasp type, as changing their grasp type would not have an impact on interaction as far as the users are concerned. This is based on the fact that users deviated away from the grasp type they were instructed to use (i.e. medium wrap) in some instances in this work, and this shows that there is still a problem for interaction designers in deciding what type of grasps users should use, particularly for natural user interfaces where there is currently no clear taxonomy that clearly defines the most suitable grasp types for interacting with virtual objects like with real objects (Feix et al., 2009). This work recommends that such a taxonomy is required for AR environments where more grasp types can be assessed and used, as assuming that physical grasping would function similarly in AR environments as it does in the real world can potentially be incorrect. Physical grasping of real objects also assumes that a power grasp reaches a definite static phase once an object is grasped (Landsmeer, 1962), however as illustrated by findings in the work of Al-Kalbani et al. (2017), a power grasp (i.e. medium wrap) does not necessarily reach this definite static phase once an object is grasped, where it was found that users generally changed their grasp type and GAp during the transition phase once the pre-load phase was complete.

9.3.3 Grasp Phases

Grasp phases are key to the planning of a grasping movement. This work mainly focused on the reaching and pre-load phases of a grasping movement, and insights regarding the final transition phase were also presented in Study 4. Findings from the four user studies in this work highlighted fundamental differences in the impact of phases on grasping movement between grasping real objects and freehand grasping of virtual objects in exocentric AR.

Separating the reaching and pre-load phases in freehand grasping was expectedly found to be challenging due to the missing physical properties of virtual objects. For example in grasping real objects, the reaching phase ends once the grasp becomes in contact with, or wraps around, the surface of the object (Gordon, 1994; Jeannerod, 1986; MacKenzie and Iberall, 1994). In freehand AR this physical indication that the reaching phase is over is absent due to the lack of tactile feedback on the hand. In studies 1-3 in this work the pre-load phase was assumed to be the moment when users verbally informed the test coordinator that they are satisfied with their grasp, this was the phase of grasping movement where users felt most confident with their grasp placement and structure, which is essentially the definition of the pre-load phase in grasping

real objects (Gordon, 1994). In Study 4, the user tolerances applied during task-based grasping interactions enforced this pre-load phase on users during the interaction, through requiring users to firstly complete the pre-load phase (i.e. trigger the interaction by satisfying the tolerances applied using their grasp) before translating the object to a target location. However even though this work recreated the pre-load phase for freehand grasping to assess accuracy during this particular phase, it is still evident that distinguishing between the two phases (reaching and pre-load) is not as straightforward in AR environments as it is for grasping real objects due to the lack of sensory information on the hand, and the lack of physical properties in virtual objects such as mass and gravity. Users across all studies in this thesis were generally in contact with the object either spatially or at least visually before completing the pre-load phase, thus the two phases were not separated automatically. This shows that transferring the assumptions associated with grasp phases for grasping real objects to freehand grasping in AR environments is not directly suitable without additional sensory feedback, and revising these phases for freehand grasping is required to make them more fitting to this natural form of grasping and easier to separate. The following grasp phases for freehand grasping that is mediated without the use of any wearable sensors are proposed:

- **Reaching:** this particular phase remains unchanged as it occurs on the same anatomical and biomechanical levels for grasping both real and virtual objects. The arm moves towards the virtual object, and the fingers are preshaped to accommodate the size and shape of the object. This phase ends when the hand is in contact with the object either spatially (i.e. hand is co-located with the virtual object in three-dimensional space) or visually (i.e. grasp looks like it is on the object in the feedback method used).

- **Grasping:** this phase replaces the pre-load phase that is highly influenced by physical task constraints that are absent in freehand grasping (e.g. weight, haptic feedback and friction). This phase starts when the hand is in contact with the object, either spatially or visually, and a certain grasp type is formed. This phase ends once the virtual object starts moving.

- **Translating:** this phase starts once the virtual object starts moving. However the assumptions associated with this particular phase in grasping real objects such as that grasping accuracy and structure remains unchanged during this phase should be addressed, as Al-Kalbani et al. (2017) showed that grasping accuracy and structure significantly change during transition of virtual objects given that the system is reliable. This phase ends once the grasping task is completed.

Revising grasp phases for natural user interfaces can potentially be useful for natural AR applications where analysing user performance in separate phases of grasping is required.

9.4 Summary

This chapter provided a discussion of findings and usability insights from all four user studies presented in this thesis (Chapters 5-8). Contributions of this work were presented, and the potential benefits of these contributions to the research community and future AR systems were also discussed. Finally this chapter also presented usability and design recommendations for future AR systems and interaction designers, and discussed the suitability and impact of physical grasping parameters such as grasp types and phases when transferred to natural AR interfaces. In Chapter 10 the conclusions drawn from the work in this thesis will be presented, along with a full summary of the findings from the user studies in Chapters 5-8.

Chapter 10

Conclusions

This thesis investigated the problem of using the natural interaction technique that is grasping in exocentric AR environments (see Table 10.1 [page 205] for a summary of findings). The primary aim of this work was to evaluate in quantifiable measures the accuracy and usability of freehand grasping in AR environments. This was achieved through four independent user studies that highlighted the key problems associated with freehand grasping. Methods for improving user performance in grasping were assessed, namely dual view visual feedback (Chapter 6), drop shadows (Chapter 7) and application of user-based tolerances (Chapter 8). Usability of freehand grasping using these methods was also evaluated using the standardised SUS test (Brooke, 1996). Knowledge of the accuracy and problems of freehand grasping can be used to aid in the development of more usable AR systems, and particularly aid in bridging the gap between reality and virtuality by allowing users to use natural interaction techniques, similar to the ones used in real life, to interact with virtual objects. In addition, the usability and accuracy of grasping can provide interaction designers with various interaction solutions through utilising the unique interplay between the fingers and potentially access the 33 different physical grasp types; this unique interplay is not necessarily present in gesture-based interaction techniques and could aid in increasing attachment and connection when interacting with virtual objects. This is particularly true in AR environments where the majority of the environment is real, thus it seems reasonable to use natural interaction techniques such as grasping in these environments where users can visualise their real hand and its interactions. Section 10.1 [page 206] will summarise the main findings in the four user studies in this thesis, and finally Section 10.2 [page 208] will highlight the limitations in this work, along with recommendations for further research.

Table 10.1: Summary of findings across the four user studies in this thesis. Methods that had a significant impact in improving freehand grasping in terms of the four measures used in this work (GAp, $GDisp_{\mathrm{x}}$, $GDisp_{\mathrm{y}}$, $GDisp_{\mathrm{z}}$, task completion time) are marked with (✓). Methods that resulted in comparable results between conditions or had no significant impact on grasping accuracy and usability are marked with (–)

Study [Method]	**No. of Grasps Assessed**	**Grasp Aperture (GAp) [mm]**	**Grasp Displacement - x axis ($GDisp_{\mathbf{x}}$) [mm]**	**Grasp Displacement - y axis ($GDisp_{\mathbf{y}}$) [mm]**	**Grasp Displacement - z axis ($GDisp_{\mathbf{z}}$) [mm]**	**Completion Time [s]**
Study 1 [Baseline]	1710	-	-	-	-	-
Study 2 [Dual view visual feedback]	1710	-	✓	✓	✓	-
Study 3 [Drop shadows]	1620	-	-	-	✓	✓
Study 4 [User based tolerances]	720	-	-	-	-	✓

Table 10.2: Summary of usability findings across the three studies in this thesis where usability was measured (1-3). Scores and labels are based on the SUS usability test (Brooke, 1996; Bangor et al., 2009*a*)

Study [Method]	No. of Grasps Assessed	Usability Score	Usability Label
Study 2 [Dual view visual feedback]	1710	Object Size: 77.00 ± 16.45	Good and Acceptable
		Object Position: 64.50 ± 13.43	Ok and Marginally Acceptable
Study 3 [Drop shadows]	1620	Drop Shadows: 81.16 ± 11.55	Good and Highly Acceptable
		No Drop Shadows: 78.17 ± 15.92	Ok and Marginally Acceptable
Study 4 [User based tolerances]	720	Absolute: 62.83 ± 18.68	Ok
		Average: 70.16 ± 14.04	Good

10.1 Review of Research

10.1.1 Measuring Accuracy of Freehand Grasping in AR

The first step in this assessment of freehand grasping was designing new metrics that can be used to assess the accuracy and usability of freehand grasping in AR. Two metrics, namely Grasp Aperture (GAp) and Grasp Displacement ($GDisp$), were introduced to quantify user performance in freehand grasping. GAp measured the Euclidean distance between the user's index finger and thumb, and is a widely used metric to assess user performance in physical grasping (Edsinger and Kemp, 2007; MacKenzie and Iberall, 1994). $GDisp$ was also introduced to be used alongside GAp, to provide information regarding the position and placement of a grasp in 3D space in relation to a virtual object. The usability of methods to assist in freehand grasping (i.e. dual view visual feedback, drop shadows and user based tolerances), was assessed by the standardised System Usability Scale (SUS) (Brooke, 1996), alongside task completion time to further assess the suitability of the three methods for accuracy (e.g. surgical training) or speed (e.g. task based rehabilitation) centred AR applications.

10.1.2 User Studies

Freehand grasping accuracy and usability were assessed in four independent user studies (120 participants and 5760 grasping tasks in total, 30 participants for each study). Table 10.1 [page 205] shows a summary of findings across the four user studies.

Study 1 (Chapter 5) presented a first analysis into freehand grasping in exocentric AR (Al-Kalbani et al., 2016*a*) that provides a comprehensive analysis of 1710 grasping tasks of virtual objects in different sizes, positions and types. This study highlighted two key problems in freehand grasping, namely significant underestimation in depth of virtual object position in the z axis and inaccurate size estimation using GAp. Depth underestimation for freehand grasping was attributed to the feedback method used being a single monitor. The assessment of GAp in relation to virtual object size was also valuable in providing insights regarding user preferences in freehand

grasping, namely that users performed a GAp that ranged from 60mm to 80mm despite being presented with object sizes that ranged from 40mm to 100mm in size.

Study 2 (Chapter 6) replicated the experiments in Study 1 (Al-Kalbani et al., 2016*b*) with the addition of dual view visual feedback in an attempt to address the two problems found in the first study (i.e. depth underestimation and inaccurate size estimation), in addition to assessing the usability of the dual view feedback method used. Findings showed that dual view visual feedback significantly improves depth estimation of virtual objects, through providing users with a side view of their interaction that enabled them to visualise their interaction in the z axis and increased their awareness of grasping errors in grasp placement or size. Dual view visual feedback also increased task completion time significantly, this is attributed to the additional side view that aided users in spending more time in correcting their grasping performance. Nevertheless dual view visual feedback was rated by users as "good and acceptable" according to the rating of Bangor et al. (2009*a*) (SUS score - 77.00) for grasping objects in different sizes, while it was rated as "OK and marginally acceptable" (SUS score - 64.50) for grasping objects in different positions (see Table 10.2 [page 206]). In the post test questionnaire, users noted that the additional side view as necessary for accurate grasping, and that it provides important information regarding the depth position of the object.

In Study 3 (Chapter 7), drop shadows were used as an additional depth cue. This study measured the impact of drop shadows when used as a depth cue on freehand grasping accuracy, and evaluated the usability of drop shadows in exocentric AR. Drop shadows significantly reduced task completion time compared to dual view visual feedback, and this was attributed to the additional information that shadows present for users even prior to starting their grasping movements. Drop shadows have also significantly improved depth estimation when compared to freehand grasping without using drop shadows. In terms of usability, users rated the use of drop shadows for freehand grasping as "GOOD and highly acceptable" (SUS score - 81.16), this was the highest usability score across all the methods proposed in this work to assist users during freehand grasping (see Table 10.2 [page 206]).

In Study 4 (Chapter 8), the user errors found in Study 1 (Chapter 5) were applied as tolerances to assist users in freehand grasping. Tolerances were applied in two configurations, namely absolute to the unique object position and average that is generalised for a z plane in which the object is located, where users were required to perform a grasp that is within both ranges of the tolerances applied (i.e GAp and $GDisp$) in order to trigger a visual indication that the grasp is acceptable and finish the task. Findings in this study have shown the application of average tolerances significantly reduces task completion time, thus offering an effective alternative solution to drop shadows and dual view visual feedback in improving task completion time. In terms of usability, users rated average tolerances as "Good" (SUS score - 70.16) and absolute tolerances as "Ok" (SUS score - 62.83) (see Table 10.2 [page 206]). Users have also indicated in the post test questionnaires that the application of tolerances during freehand grasping of virtual objects is easy to learn. This study showed that freehand grasping can be usable to com-

plete interaction tasks without being required to be highly accurate in virtual object site and position.

Developing AR systems that allow users to interact with virtual objects with a usable natural form of grasping (i.e. freehand grasping) can potentially aid novice users of AR technology to to be more connected to the technology through this natural interaction form in an accurate and usable manner as demonstrated by the different methods in this thesis. Furthermore freehand grasping is particularly important for applications where the use of additional wearable devices for haptic feedback is not desirable, due to their discomfort for users (Kimura and Sato, 2018) or restrictive setup and configuration (Bikos et al., 2015). The human hand is a strong tool that is widely used on daily basis, natural freehand interaction would ease the process of training novice users to AR technology, potentially increasing the social and individual acceptance of AR through bridging the gap between virtuality and reality when users are presented with elements of both worlds simultaneously in AR environments. This work provides evidence that natural grasping can be usable in AR applications and tasks, alongside methods that can significantly improve natural grasping performance.

10.2 Constraints and Future Work

10.2.1 Environment

This first assessment of freehand grasping was implemented in an exocentric AR environment. Users in this work viewed the environment from the outside using a large feedback monitor displaying a composited real-time mirrored scene, overlaying virtual objects with the video feed, this can be comparable to CAVE (e.g. Liu and Cheung (2016)) and projection based (e.g. Besharati Tabrizi and Mahvash (2015)) AR systems, and is particularly useful when the use of wearable devices is undesirable due to their restrictive setup or when additional sensors are not practical or feasible to use in certain domains such as medical AR applications. While this environment was suitable to recreate and evaluate natural grasping of virtual objects, evaluating the accuracy and usability of grasping in one user interface is still limiting. Future work should translate and validate current findings in other user interfaces that do not necessarily use a single view as a feedback method. Assessing freehand grasping in egocentric AR environments should also be considered in future work owing to the wide use of HMDs in current research to mediate hand-based interaction with virtual objects. AR egocentric environments allow users to fully visualise the depth and three-dimensional position of virtual objects, this can potentially have a significant impact on freehand grasping accuracy in AR.

In addition, only abstract or regular object types (i.e. cubes and spheres) were used in this first assessment of freehand grasping. This was a control measure for this first assessment of grasping in exocentric AR. Complex objects in terms of structure and rendering quality should be considered in future work. Presenting users with complex objects such as objects with handles

(e.g. mugs) or floppy objects (e.g. sponges) can further assess the usability of natural grasping of virtual objects (Feix et al., 2009), and also aid in better understanding how user grasp strategy changes when presented with complex virtual objects that are similar to objects that are often grasped in real environments. This work assessed the use of drop shadows to assist in freehand grasping, where drop shadows significantly improved usability by providing users with an additional visual cue during the interaction. Owing to the fact that visual perception of objects is a major part of freehand grasping with the absence of tactile feedback, other rendering techniques such as shading, textures and lighting should be implemented in future work to assess their impact on freehand grasping.

10.2.2 Methods and Transferability

This thesis presented various methods for assessing and improving freehand grasping in terms of user accuracy and usability using the SUS test, these methods can be further developed in different user interfaces. For example, dual view visual feedback has shown in this work that it is an effective method for improving virtual object depth estimation. However, the impact or feasibility of this method in wearable based AR applications is not yet clear. Future work should build on previous research that focused on multiple viewpoints for AR based interaction (Hoang et al., 2011) and translate this method to egocentric AR environments where depth underestimation is still problematic.

Furthermore, user-based tolerances in this work significantly improved usability in freehand grasping. However tolerances in this work were fixed values that were essentially user errors in Study 1 (Chapter 5), this can be limiting when the task requirements in a specific system change. For example, if a certain task requires users to be accurate in grasping, fixed values may not be effective in assisting users to be accurate and vice versa for speed. Future work can further develop tolerances to be automated based on task requirements, thus if a task requires users to be fast during grasping then tolerances should automatically be adjusted to aid users in achieving that speed in interaction.

In addition, future work should also investigate how combining the methods introduced in this thesis can impact freehand grasping performance and usability. For example, users in this work suggested in the post test questionnaires that dual view visual feedback and drop shadows could potentially be used together. Based on the findings from the two studies that presented drop shadows and dual view visual feedback, combining the two methods can potentially improve the usability of the dual view visual feedback method and also potentially lead to higher accuracy in freehand grasping.

Finally, this work highlighted that size estimation in freehand grasping is significantly problematic. Even though the application of user-based tolerances negated this problem by allowing users to complete grasping tasks without the need to be accurate in size and position estimation, users were highly inaccurate in estimating virtual object size using GAp in the three studies that assessed grasping accuracy and usability (studies 1, 2 and 3). Accurate size estimation

is highly influenced by the presence of tactile feedback on the hand, and future work should investigate how wearable based grasping with haptic feedback on the hand can improve grasping, and how it compares in terms of accuracy and usability to freehand grasping that was assessed in this work. Haptic feedback can potentially mitigate this problem in size estimation. Furthermore investigating the impact of haptic feedback can also illustrate whether compensating tactile feedback is enough for grasping virtual objects to be as accurate as the physical grasping of real objects or at least more accurate than freehand grasping, or whether other parameters such as object weight and friction are more prominent factors for achieving accurate grasping. Current research is also exploring how audio feedback can be used as a feedback method for freehand grasping interactions (Kimura and Sato, 2018), further work can explore how audio feedback can impact size estimation in AR applications.

10.2.3 Interaction Technique

This work presented a first analysis of the accuracy and usability of freehand grasping in exocentric AR. For this reason, only one grasp type was assessed in controlled experiment setups in the four user studies in this work. Using one grasp type was effective in assessing accuracy and usability of freehand grasping in this work. However, future work should assess different grasp types and validate findings presented in this thesis for other types of grasps that are widely used in real environments. Assessing more grasp types will form a better understanding of this form of natural grasping of virtual objects. In addition, the medium wrap grasp was used based on research that is focused on physical grasping of real objects, this is mainly due to the lack of grasp classification and guidelines for natural grasping in AR environments. Findings in this thesis showed that users deviated away from the medium wrap grasp chosen for this analysis, thus indicating that grasp choice is potentially formulated by users using different parameters and design choices when naturally grasping virtual objects to the ones outlined for grasping real objects. This shows that future work should also focus on forming a taxonomy for this natural form of grasping virtual objects. Through understanding the factors that impact grasp choice and strategy, a grasp taxonomy will aid interaction designers in using suitable grasp types for different applications in AR environments.

Appendix A

Post-Test Questionnaires - Dual View Visual Feedback (Chapter 6)

GRASPING VIRTUAL OBJECTS USING MULTI-VIEW FEEDBACK – POST-TEST SURVEY

User Code: ……

PART 1 – *Please tick or highlight one answer (grey boxes offer further explanation of the questions)*

Based on the System Usability Scale (SUS) - © Digital Equipment Corporation 1986

1. I think that I would like to use this system frequently

If given the chance to use the system again, you would like to use the system with two feedback cameras and not just the one from the Kinect

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

2. I found the system unnecessarily complex

You think that adding a second camera makes using the system more confusing and challenging, rather than simplify its use

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

3. I thought the system was easy to use

You thought adding a second camera is easy and logical to use, and it helped in improving your grasping performance

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

4. I think that I would need the support of a technical person to be able to use this system

In order to understand how the system works with the two cameras, I needed a technical person to explain the system to me, and I required training to fully understand its functionality

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

5. I found the various functions in this system were well integrated

I found that all the components of the system (positions of 3D objects, frontal view Kinect camera, side view camera and test instructions) were well integrated and functioning

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

6. I thought there was too much inconsistency in this system

I thought that adding a new side view camera added made my performance inconsistent and prolonged my task completion time

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

7. I would imagine that most people would learn to use this system very quickly

I think that learning this system with the two cameras will be easy for anyone to learn with short training sessions

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

8. I found the system very cumbersome to use

I thought using the system with two feedback views distracting and challenging to use

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

9. I felt very confident using the system

I felt using two feedback views was easy and I was confident in my grasping performance

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

10. I needed to learn a lot of things before I could get going with this system

Training to learn how the system works with the two feedback views was long and it took a long time for me to learn how to use the system

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

PART TWO – *Please tick or highlight one answer, and then give your reasons or more details in the grey box if available*

---

1. Which screen did you look at first?

- ○ Frontal view (Kinect)
- ○ Side view (Webcam)

……

……

---

2. Which screen did you depend on the most?

- ○ Frontal view (Kinect)
- ○ Side view (Webcam)

……

……

---

3. Which view did you find to be more important?

- ○ Frontal view (Kinect)
- ○ Side view (Webcam)

……

……

---

4. Did you use the two view in a specific order? (If yes, which one? (Write it down on the comments please)

- ○ Yes
- ○ No

……

……

5. Do you think changing positions of both feedback screens would make a difference in performance? (E.g. place the side view to your right and the Kinect view to your left – or vice versa)?

- ○ Yes
- ○ No

.....

.....

General Comments

.....

.....

.....

.....

Appendix B

Post-Test Questionnaires - Drop Shadows (Chapter 7)

GRASPING VIRTUAL OBJECTS USING DROP SHADOWS – POST-TEST SURVEY

User Code: ……

PART 1 – *Please tick or highlight one answer (grey boxes offer further explanation of the questions)*

Based on the System Usability Scale (SUS) - © Digital Equipment Corporation 1986

1. I think that I would like to use this system frequently

If given the chance to use the system again, you would like to use the system with two feedback cameras and not just the one from the Kinect

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

2. I found the system unnecessarily complex

You think that adding a second camera makes using the system more confusing and challenging, rather than simplify its use

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

3. I thought the system was easy to use

You thought adding a second camera is easy and logical to use, and it helped in improving your grasping performance

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

4. I think that I would need the support of a technical person to be able to use this system

In order to understand how the system works with the two cameras, I needed a technical person to explain the system to me, and I required training to fully understand its functionality

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

5. I found the various functions in this system were well integrated

I found that all the components of the system (positions of 3D objects, frontal view Kinect camera, side view camera and test instructions) were well integrated and functioning

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

6. I thought there was too much inconsistency in this system

I thought that adding a new side view camera added made my performance inconsistent and prolonged my task completion time

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

7. I would imagine that most people would learn to use this system very quickly

I think that learning this system with the two cameras will be easy for anyone to learn with short training sessions

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

8. I found the system very cumbersome to use

I thought using the system with two feedback views distracting and challenging to use

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

9. I felt very confident using the system

I felt using two feedback views was easy and I was confident in my grasping performance

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

10. I needed to learn a lot of things before I could get going with this system

Training to learn how the system works with the two feedback views was long and it took a long time for me to learn how to use the system

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

PART TWO – *Please tick or highlight one answer, and then give your reasons or more details in the grey box if available*

---

1. I found it easy to locate and successfully grasp objects

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

2. I have noticed that the virtual objects changed position in the x, y and z axes

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

3. I have noticed the drop shadows changed in position in the x, y and z axes depending on the object's position *[drop shadows condition]*

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

4. I used the drop shadows to locate the virtual objects presented *[drop shadows condition]*

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

5. I found the drop shadows useful in accurately locating virtual objects *[drop shadows condition]*

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

6. Which depth cue did you find to be more useful in locating virtual objects? *[drop shadows condition]*

- ○ Occlusion
- ○ Drop Shadows
- ○ Both

---

7. In your opinion, would adding/enhancing these rendering features improve grasping virtual objects? *[drop shadows condition]*

- Shadows

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

- Object Lighting

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

- Object Texture

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

- Object Shape

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

- Object Size

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

8. Did you suffer from any fatigue/pain during any of the tasks in this test? If yes, please specify which tasks caused you discomfort (e.g. positions that were low, closer to screen, furthest from screen) in the grey box below *[drop shadows condition]*

.....

.....

.....

.....

---

9. Which of the two objects did you find easier to interact with?

○ Cube
○ Sphere

General Comments

.....
.....
.....
.....

Appendix C

Post-Test Questionnaires - User Based Tolerances (Chapter 8)

GRASPING AND MOVING VIRTUAL OBJECTS – POST-TEST SURVEY

User Code: ……

PART 1 – *Please tick or highlight one answer (grey boxes offer further explanation of the questions)*

Based on the System Usability Scale (SUS) - © Digital Equipment Corporation 1986

1. I think that I would like to use this system frequently

If given the chance to use the system again, you would like to use the system with two feedback cameras and not just the one from the Kinect

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

2. I found the system unnecessarily complex

You think that adding a second camera makes using the system more confusing and challenging, rather than simplify its use

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

3. I thought the system was easy to use

You thought adding a second camera is easy and logical to use, and it helped in improving your grasping performance

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

4. I think that I would need the support of a technical person to be able to use this system

In order to understand how the system works with the two cameras, I needed a technical person to explain the system to me, and I required training to fully understand its functionality

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

5. I found the various functions in this system were well integrated

I found that all the components of the system (positions of 3D objects, frontal view Kinect camera, side view camera and test instructions) were well integrated and functioning

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

6. I thought there was too much inconsistency in this system

I thought that adding a new side view camera added made my performance inconsistent and prolonged my task completion time

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

7. I would imagine that most people would learn to use this system very quickly

I think that learning this system with the two cameras will be easy for anyone to learn with short training sessions

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

8. I found the system very cumbersome to use

I thought using the system with two feedback views distracting and challenging to use

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

9. I felt very confident using the system

I felt using two feedback views was easy and I was confident in my grasping performance

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

10. I needed to learn a lot of things before I could get going with this system

Training to learn how the system works with the two feedback views was long and it took a long time for me to learn how to use the system

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

PART TWO – *Please tick or highlight one answer, and then give your reasons or more details in the grey box if available*

---

1. I found it easy to locate and successfully grasp objects

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

2. I found it easy to move objects to the target location

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

3. I have noticed the virtual objects changed position in the x, y and z axes

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

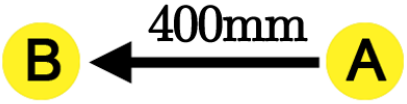
---

4. Rate the difficulty of each task you have completed

- Task 1

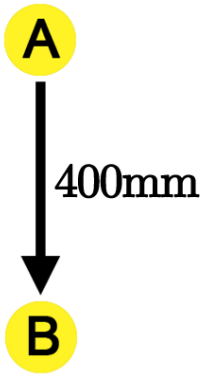
	1	2	3	4	5	
Very Difficult	○	○	○	○	○	**Very Easy**

- Task 2



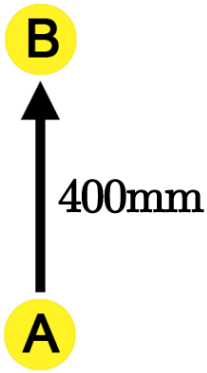
	1	2	3	4	5	
Very Difficult	○	○	○	○	○	**Very Easy**

- Task 3



	1	2	3	4	5	
Very Difficult	○	○	○	○	○	**Very Easy**

- Task 4



	1	2	3	4	5	
Very Difficult	○	○	○	○	○	**Very Easy**

---

5. In your opinion, would adding/enhancing these rendering features improve grasping virtual objects?

- Shadows

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

- Object Lighting

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

- Object Texture

- Object Shape

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

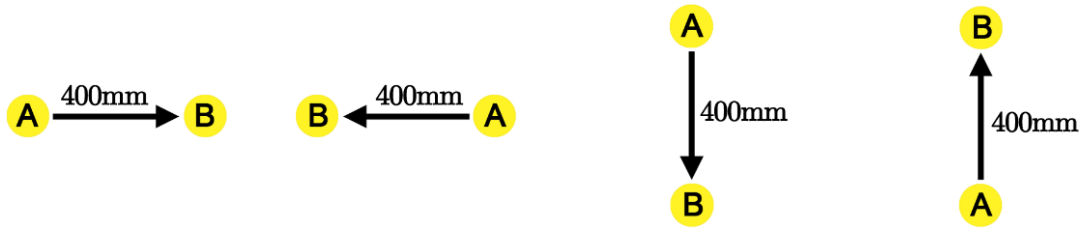
- Object Size

	1	2	3	4	5	
Totally Disagree	○	○	○	○	○	Totally Agree

---

6. Did you suffer from any fatigue/pain during any of the tasks in this test? If yes, please specify tick below task(s) that caused you discomfort

○ Yes
○ No



TASK 1 ○ TASK 2 ○ TASK 3 ○ TASK 4 ○

---

7. Which of the two objects did you find easier to interact with?

○ Cube
○ Sphere

General Comments

.....

.....

.....

.....

Bibliography

Al-Kalbani, M., Frutos-Pascual, M. and Williams, I. (2017), Freehand grasping in mixed reality: analysing variation during transition phase of interaction, *in* 'Proceedings of the 19th ACM International Conference on Multimodal Interaction', ACM, pp. 110–114. Cited on pages 6, 172, 174, 201, and 202.

Al-Kalbani, M., Williams, I. and Frutos-Pascual, M. (2016*a*), Analysis of medium wrap freehand virtual object grasping in exocentric mixed reality, *in* 'Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on', IEEE, pp. 84–93. Cited on pages 6, 66, and 206.

Al-Kalbani, M., Williams, I. and Frutos-Pascual, M. (2016*b*), Improving freehand placement for grasping virtual objects via dual view visual feedback in mixed reality, *in* 'Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology', ACM, pp. 279–282. Cited on pages 6, 98, and 207.

Albert, W. and Tullis, T. (2013), *Measuring the user experience: collecting, analyzing, and presenting usability metrics*, Newnes. Cited on page 184.

Aleotti, J. and Caselli, S. (2006), Grasp recognition in virtual reality for robot pregrasp planning by demonstration, *in* 'Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on', IEEE, pp. 2801–2806. Cited on page 41.

Allen, G. and Tsukahara, N. (1974), 'Cerebrocerebellar communication systems.', *Physiological reviews* **54**(4), 957–1006. Cited on page 33.

Anderson, J. R. (1985), *Cognitive psychology and its implications*, WH Freeman/Times Books/Henry Holt & Co. Cited on pages 32, 44, and 190.

Arbib, M., Iberall, T., Lyons, D. and Linscheid, R. (1985), 'Hand function and the neocortex', *Springer-Verlag, ch. Coordinated control programs for movement of the hand* pp. 111–129. Cited on page 38.

Arkenbout, E. A., de Winter, J. C. and Breedveld, P. (2015), 'Robust hand motion tracking through data fusion of 5dt data glove and nimble vr kinect camera measurements', *Sensors* **15**(12), 31644–31671. Cited on page 24.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S. and MacIntyre, B. (2001), 'Recent advances in augmented reality', *IEEE computer graphics and applications* **21**(6), 34–47. Cited on page 12.

Bai, H., Gao, L., El-Sana, J. and Billinghurst, M. (2013), 'Markerless 3D gesture-based interaction for handheld Augmented Reality interfaces', *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013* (October 2013), 0–5. Cited on pages 19 and 20.

Bangor, A., Kortum, P. and Miller, J. (2009*a*), 'Determining what individual sus scores mean: Adding an adjective rating scale', *Journal of usability studies* **4**(3), 114–123. Cited on pages XVIII, 112, 134, 168, 197, 206, and 207.

Bangor, A., Kortum, P. and Miller, J. (2009*b*), 'Determining what individual sus scores mean: Adding an adjective rating scale', *Journal of usability studies* **4**(3), 114–123. Cited on pages 181 and 191.

Benko, H., Jota, R. and Wilson, A. (2012), 'MirageTable: Freehand Interaction on a Projected Augmented Reality Tabletop', *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* pp. 199–208. Cited on pages 18, 19, and 25.

Bennett, K. M. and Castiello, U. (1994), *Insights into the reach to grasp movement*, Vol. 105, Elsevier. Cited on pages 31, 33, and 45.

Besharati Tabrizi, L. and Mahvash, M. (2015), 'Augmented reality–guided neurosurgery: accuracy and intraoperative application of an image projection technique', *Journal of neurosurgery* **123**(1), 206–211. Cited on page 208.

Bikos, M., Itoh, Y., Klinker, G. and Moustakas, K. (2015), An interactive augmented reality chess game using bare-hand pinch gestures, *in* 'Cyberworlds (CW), 2015 International Conference on', IEEE, pp. 355–358. Cited on page 208.

Billinghurst, M. and Buxton, B. (2011), 'Gesture based interaction', *Haptic input* **24**. Cited on page 25.

Billinghurst, M., Clark, A., Lee, G. et al. (2015), 'A survey of augmented reality', *Foundations and Trends® in Human–Computer Interaction* **8**(2-3), 73–272. Cited on pages 13 and 193.

Billinghurst, M. and Kato, H. (1999), Collaborative mixed reality, *in* 'Proceedings of the First International Symposium on Mixed Reality', pp. 261–284. Cited on page 10.

Billinghurst, M., Piumsomboon, T. and Bai, H. (2014), 'Hands in space: Gesture interaction with augmented-reality interfaces', *IEEE computer graphics and applications* **34**(1), 77–80. Cited on pages 19 and 20.

Blaga, A.-D., Frutos-Pascual, M., Al-Kalbani, M. and Williams, I. (2017), [poster] usability analysis of an off-the-shelf hand posture estimation sensor for freehand physical interaction in

egocentric mixed reality, *in* 'Mixed and Augmented Reality (ISMAR-Adjunct), 2017 IEEE International Symposium on', IEEE, pp. 31–34. Cited on page 26.

Bock, O. and Jüngling, S. (1999), 'Reprogramming of grip aperture in a double-step virtual grasping paradigm', *Experimental brain research* **125**(1), 61–66. Cited on pages 15 and 188.

Borst, C. W. and Indugula, A. P. (2006), 'A spring model for whole-hand virtual grasping', *Presence: Teleoperators and Virtual Environments* **15**(1), 47–61. Cited on page 16.

Borst, C. W. and Prachyabrued, M. (2013), Nonuniform and adaptive coupling stiffness for virtual grasping, *in* 'Virtual Reality (VR), 2013 IEEE', IEEE, pp. 35–38. Cited on pages 16 and 17.

Bowman, D. A. (1999), 'Interaction techniques for common tasks in immersive virtual environments', *Georgia Institute of Technology* . Cited on pages 11 and 34.

Bozzacchi, C. and Domini, F. (2015), 'Lack of depth constancy for grasping movements in both virtual and real environments', *Journal of neurophysiology* **114**(4), 2242–2248. Cited on page 189.

Brooke, J. (1996), 'Sus-a quick and dirty usability scale', *Usability evaluation in industry* **189**(194), 4–7. Cited on pages XVIII, 103, 142, 204, and 206.

Brooke, J. et al. (1996), 'Sus-a quick and dirty usability scale', *Usability evaluation in industry* **189**(194), 4–7. Cited on pages 180 and 181.

Bruder, G., Sanz, F. A., Olivier, A.-H. and Lécuyer, A. (2015), Distance estimation in large immersive projection systems, revisited, *in* 'Virtual Reality (VR), 2015 IEEE', IEEE, pp. 27–32. Cited on page 193.

Buchmann, V., Violich, S., Billinghurst, M. and Cockburn, A. (2004), 'FingARtips: gesture based direct manipulation in Augmented Reality', *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia* pp. 212–221. Cited on pages IX, 17, 18, and 24.

Bullock, I. M., Feix, T. and Dollar, A. M. (2013), 'Finding small, versatile sets of human grasps to span common objects', *Proceedings - IEEE International Conference on Robotics and Automation* pp. 1068–1075. Cited on pages X, 39, and 49.

Cai, M., Kitani, K. M. and Sato, Y. (2017), 'An ego-vision system for hand grasp analysis', *IEEE Transactions on Human-Machine Systems* . Cited on page 41.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E. and Ivkovic, M. (2011), 'Augmented reality technologies, systems and applications', *Multimedia Tools and Applications* **51**(1), 341–377. Cited on page 12.

Castiello, U. (2005), 'The neuroscience of grasping', *Nature Reviews Neuroscience* **6**(9), 726–736. Cited on page 42.

Chan, T.-c., Carello, C. and Turvey, M. (1990), 'Perceiving object width by grasping', *Ecological Psychology* **2**(1), 1–35. Cited on page 59.

Chapin, J. K. and Moxon, K. A. (2000), *Neural prostheses for restoration of sensory and motor function*, CRC Press. Cited on pages 31 and 32.

Chen, K. B., Kimmel, R. A., Bartholomew, A., Ponto, K., Gleicher, M. L. and Radwin, R. G. (2014), 'Manually locating physical and virtual reality objects', *Human factors* **56**(6), 1163–1176. Cited on pages 198 and 199.

Chen, Z. and Saunders, J. A. (2016), 'Automatic adjustments toward unseen visual targets during grasping movements', *Experimental brain research* **234**(7), 2091–2103. Cited on pages XV, 192, and 193.

Choi, I., Hawkes, E. W., Christensen, D. L., Ploch, C. J. and Follmer, S. (2016), Wolverine: A wearable haptic interface for grasping in virtual reality, *in* 'Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on', IEEE, pp. 986–993. Cited on pages 16 and 17.

Cidota, M. A., Clifford, R. M. S., Dezentje, P., Lukosch, S. G. and Bank, P. J. M. (2015), 'Affording Visual Feedback for Natural Hand Interaction in AR to Assess Upper Extremity Motor Dysfunction', *2015 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* pp. 1–4. Cited on pages 21, 22, and 25.

Cidota, M. A., Lukosch, S. G., Dezentje, P., Bank, P. J., Lukosch, H. K. and Clifford, R. (2016), 'Serious gaming in augmented reality using hmds for assessment of upper extremity motor dysfunctions', *i-com* **15**(2), 155–169. Cited on pages 2, 21, 23, and 25.

Corera, S. and Krishnarajah, N. (2011), 'Capturing hand gesture movement: a survey on tools, techniques and logical considerations', *Proceedings of chi sparks* . Cited on page 23.

Cutkosky, M. R. (2012), *Robotic grasping and fine manipulation*, Vol. 6, Springer Science & Business Media. Cited on pages 31 and 32.

Cutkosky, M. R. and Howe, R. D. (1990), 'Human Grasp Choice and Robotic Grasp Analysis', *Dextrous Robot Hands* **1**, 5–31. Cited on pages 34, 37, and 40.

Cutkosky, M. and Wright, P. (1986), Modeling manufacturing grips and correlations with the design of robotic hands, *in* 'Robotics and Automation. Proceedings. 1986 IEEE International Conference on', Vol. 3, IEEE, pp. 1533–1539. Cited on pages 37 and 40.

Cutting, J. E. (1997), 'How the eye measures reality and virtual reality', *Behavior Research Methods, Instruments, & Computers* **29**(1), 27–36. Cited on page 61.

Datcu, D. and Lukosch, S. (2013), 'Free-hands interaction in augmented reality', *Proceedings of the 1st symposium on Spatial user interaction - SUI '13* pp. 33–40. Cited on pages 20, 21, and 25.

Davis, M. M., Gabbard, J. L., Bowman, D. A. and Gracanin, D. (2016), Depth-based 3d gesture multi-level radial menu for virtual object manipulation, *in* 'Virtual Reality (VR), 2016 IEEE', IEEE, pp. 169–170. Cited on pages 20 and 21.

de La Gorce, M., Paragios, N. and Fleet, D. J. (2008), Model-based hand tracking with texture, shading and self-occlusions, *in* 'Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference On', IEEE, pp. 1–8. Cited on page 25.

Debowy, D. J., Ghosh, S., Ro, J. Y. and Gardner, E. P. (2001), 'Comparison of neuronal firing rates in somatosensory and posterior parietal cortex during prehension', *Experimental Brain Research* **137**(3), 269–291. Cited on page 44.

Diaz, C., Walker, M., Szafir, D. A. and Szafir, D. (2017), Designing for depth perceptions in augmented reality, *in* '2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)', IEEE, pp. 111–122. Cited on pages 138 and 196.

Drascic, D. and Grodski, J. (1993), 'Using stereoscopic video for bomb disposal teleoperation', *SPIE 1915: Stereoscopic Displays and Applications IV* . Cited on page 193.

Drascic, D. and Milgram, P. (1996), Perceptual issues in augmented reality, *in* 'Proceedings-Spie The International Society For Optical Engineering', SPIE INTERNATIONAL SOCIETY FOR OPTICAL, pp. 123–134. Cited on pages 193 and 194.

Duff, M., Chen, Y., Attygalle, S., Herman, J., Sundaram, H., Qian, G., He, J. and Rikakis, T. (2010), 'An adaptive mixed reality training system for stroke rehabilitation', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **18**(5), 531–541. Cited on page 99.

Dunn, O. J. (1961), 'Multiple Comparisons Among Means', *Journal of the American Statistical Association* **56**(293), 52–64. Cited on pages 68 and 103.

Dünser, A. and Billinghurst, M. (2011), Evaluating augmented reality systems, *in* 'Handbook of augmented reality', Springer, pp. 289–307. Cited on page 47.

Dünser, A., Grasset, R. and Billinghurst, M. (2008), *A survey of evaluation techniques used in augmented reality studies*, Human Interface Technology Laboratory New Zealand. Cited on page 47.

Dünser, A., Grasset, R., Seichter, H. and Billinghurst, M. (2007), 'Applying hci principles to ar systems design'. Cited on page 22.

Edsinger, A. and Kemp, C. C. (2007), Human-robot interaction for cooperative manipulation: Handing objects to one another, *in* 'Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on', IEEE, pp. 1167–1172. Cited on pages 50, 51, 188, and 206.

Ekvall, S. and Kragic, D. (2005), Grasp recognition for programming by demonstration, *in* 'Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on', IEEE, pp. 748–753. Cited on page 41.

Epstein, W. and Rogers, S. (1995), *Perception of space and motion*, Elsevier. Cited on page 61.

Erol, A., Bebis, G., Nicolescu, M., Boyle, R. D. and Twombly, X. (2007), 'Vision-based hand pose estimation: A review', *Computer Vision and Image Understanding* **108**(1), 52–73. Cited on page 27.

Evans, G., Miller, J., Pena, M. I., MacAllister, A. and Winer, E. (2017), Evaluating the microsoft hololens through an augmented reality assembly application, *in* 'Degraded Environments: Sensing, Processing, and Display 2017', Vol. 10197, International Society for Optics and Photonics, p. 101970V. Cited on pages 2, 175, and 187.

Feix, T., Bullock, I. M. and Dollar, A. M. (2014), 'Analysis of Human Grasping Behavior: Correlating Tasks, Objects and Grasps', *IEEE Transactions on Haptics (in press)* **7**(3), 311–323. Cited on pages 37, 39, 49, 56, and 72.

Feix, T., Pawlik, R., Schmiedmayer, H. B., Romero, J. and Kragi, D. (2009), 'A comprehensive grasp taxonomy', *Robotics, Science and Systems Conference: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation* pp. 2–3. Cited on pages 30, 34, 37, 39, 41, 49, 201, and 209.

Feix, T., Romero, J., Schmiedmayer, H.-B., Dollar, A. M. and Kragic, D. (2016), 'The grasp taxonomy of human grasp types', *IEEE Transactions on Human-Machine Systems* **46**(1), 66–77. Cited on pages 39 and 49.

Field, A. (2012), *Discovering statistics using R*, Sage publications. Cited on page 68.

Frajhof, L., Borges, J., Hoffmann, E., Lopes, J. and Haddad, R. (2018), 'Virtual reality, mixed reality and augmented reality in surgical planning for video or robotically assisted thoracoscopic anatomic resections for treatment of lung cancer', *The Journal of Visualized Surgery* **4**(7). Cited on page 175.

Gavrilova, M. L., Wang, Y., Ahmed, F. and Paul, P. P. (2018), 'Kinect sensor gesture and activity recognition: New applications for consumer cognitive systems', *IEEE Consumer Electronics Magazine* **7**(1), 88–94. Cited on page 58.

Gibson, J. J. (1950), 'The perception of the visual world.'. Cited on page 60.

Gibson, J. J. (1966), 'The senses considered as perceptual systems.'. Cited on page 60.

Gibson, J. J. (2014), *The ecological approach to visual perception: classic edition*, Psychology Press. Cited on page 60.

Gigante, M. A. (1993), 'Virtual reality: definitions, history and applications', *Virtual Reality Systems* pp. 3–14. Cited on page 10.

Gonzalez-Jorge, H., Rodríguez-Gonzálvez, P., Martínez-Sánchez, J., González-Aguilera, D., Arias, P., Gesto, M. and Díaz-Vilariño, L. (2015), 'Metrological comparison between kinect i and kinect ii sensors', *Measurement* **70**, 21–26. Cited on page 58.

Gordon, A. (1994), 'Development of the reach to grasp movement', *Advances in psychology* **105**, 37–56. Cited on pages 44, 50, 201, and 202.

Grinshpoon, A., Sadri, S., Loeb, G. J., Elvezio, C., Siu, S. and Feiner, S. K. (2018), Hands-free augmented reality for vascular interventions, *in* 'ACM 2018 SIGGRAPH Emerging Technologies', SIGGRAPH '18, ACM, pp. 8a:1–8a:2. Cited on page 171.

Ha, T., Feiner, S. and Woo, W. (2014), Wearhand: Head-worn, rgb-d camera-based, bare-hand user interface with visually enhanced depth perception, *in* 'IEEE ISMAR', IEEE, pp. 219–228. Cited on page 3.

Hasenfratz, J.-M., Lapierre, M., Holzschuch, N. and Sillion, F. (2003), A survey of real-time soft shadows algorithms, *in* 'Computer Graphics Forum', Vol. 22, Wiley Online Library, pp. 753–774. Cited on page 138.

Hatscher, B., Luz, M., Nacke, L. E., Elkmann, N., Müller, V. and Hansen, C. (2017), Gazetap: towards hands-free interaction in the operating room, *in* 'Proceedings of the 19th ACM International Conference on Multimodal Interaction', ACM, pp. 243–251. Cited on page 14.

Helander, M. G. (2014), *Handbook of human-computer interaction*, Elsevier. Cited on pages 32 and 47.

Hix, D. and Hartson, H. R. (1993), *Developing user interfaces: ensuring usability through product & process*, John Wiley & Sons, Inc. Cited on pages 8 and 11.

Hoang, T., Thomas, B. H. et al. (2011), Multiple camera augmented viewport: an investigation of camera position, visualizations, and the effects of sensor errors and head movement, PhD thesis, Virtual Reality Society of Japan. Cited on pages 195 and 209.

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**(3), 33–33. Cited on page 194.

Holz, D., Ullrich, S., Wolter, M. and Kuhlen, T. (2008), 'Multi-Contact Grasp Interaction for Virtual Environments', *Journal of Virtual Reality and Broadcasting* **5**(7), 101–112. Cited on pages 8 and 23.

Hondori, H. M., Khademi, M., Dodakian, L., Cramer, S. C. and Lopes, C. V. (2013), A spatial augmented reality rehab system for post-stroke hand rehabilitation., *in* 'MMVR', pp. 279–285. Cited on pages 18, 19, 23, 25, and 28.

Hough, G., Williams, I. and Athwal, C. (2015), 'Fidelity and plausibility of bimanual interaction in mixed reality', *Visualization and Computer Graphics, IEEE Transactions on* **21**(12), 1377–1389. Cited on pages 10, 26, 58, 60, 61, 66, 173, 191, and 199.

Iberall, T., Bingham, G. and Arbib, M. (1986), 'Opposition space as a structuring concept for the analysis of skilled hand movements', *Experimental brain research series* **15**, 158–173. Cited on page 38.

Iberall, T., Jackson, J., Labbe, L. and Zampano, R. (1988), Knowledge-based prehension: Capturing human dexterity, *in* 'Robotics and Automation, 1988. Proceedings., 1988 IEEE International Conference on', IEEE, pp. 82–87. Cited on page 40.

Iberall, T. and MacKenzie, C. L. (1990), Opposition space and human prehension, *in* 'Dextrous robot hands', Springer, pp. 32–54. Cited on pages XVI, 42, and 43.

Inoue, D., Cho, B., Mori, M., Kikkawa, Y., Amano, T., Nakamizo, A., Yoshimoto, K., Mizoguchi, M., Tomikawa, M., Hong, J. et al. (2013), 'Preliminary study on the clinical application of augmented reality neuronavigation', *Journal of Neurological Surgery Part A: Central European Neurosurgery* **74**(02), 071–076. Cited on page 2.

Jacobson, C. and Sperling, L. (1976), 'Classification of the hand-grip: A preliminary study.', *Journal of Occupational and Environmental Medicine* **18**(6), 395–398. Cited on page 40.

Jacobson, J. and Werner, S. (2004), 'Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art', *Perception* **33**(11), 1369–1383. Cited on page 139.

Jeannerod, M. (1981), 'Intersegmental coordination during reaching at natural visual objects', *Attention and performance IX* pp. 153–168. Cited on page 44.

Jeannerod, M. (1984), 'The timing of natural prehension movements', *Journal of motor behavior* **16**(3), 235–254. Cited on page 43.

Jeannerod, M. (1986), 'The formation of finger grip during prehension. a cortically mediated visuomotor pattern', *Behavioural brain research* **19**(2), 99–116. Cited on pages 44, 50, and 201.

Jeannerod, M. (2006), *Motor cognition: What actions tell the self*, number 42, Oxford University Press. Cited on pages 33 and 42.

Jeannerod, M., Decety, J. and Goodale, M. (1990), 'Vision and action: The control of grasping'. Cited on pages 59 and 62.

Johansson, R. and Westling, G. (1984), 'Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects', *Experimental brain research* **56**(3), 550–564. Cited on page 44.

Jonckheere, A. R. (1954), 'A distribution-free k-sample test against ordered alternatives', *Biometrika* **41**(1/2), 133–145. Cited on page 68.

Jung, W. and Woo, W. T. (2017), Duplication based distance-free freehand virtual object manipulation, *in* 'Ubiquitous Virtual Reality (ISUVR), 2017 International Symposium on', IEEE, pp. 10–13. Cited on pages 22 and 24.

Kalaska, J., Caminiti, R. and Georgopoulos, A. (1983), 'Cortical mechanisms related to the direction of two-dimensional arm movements: relations in parietal area 5 and comparison with motor cortex', *Experimental Brain Research* **51**(2), 247–260. Cited on page 33.

Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F. and Miura, Y. (1980), 'Patterns of static prehension in normal hands', *American Journal of Occupational Therapy* **34**(7), 437–445. Cited on page 36.

Kapandji, I. A. (1974), *The physiology of the joints. volume 2: Lower limb*, Churchill Livingstone. Cited on page 37.

Kelly, D., Kandel, E., Schwartz, J. et al. (1985), 'Principles of neural science'. Cited on pages X and 34.

Kerous, B. and Liarokapis, F. (2016), Brain-computer interfaces-a survey on interactive virtual environments, *in* '2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)', IEEE, pp. 1–4. Cited on page 12.

Kerous, B. and Liarokapis, F. (2017), Brainchat-a collaborative augmented reality brain interface for message communication, *in* '2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)', IEEE, pp. 279–283. Cited on pages 12 and 14.

Kersten, D. and Legge, G. E. (1983), 'Convergence accommodation', *JOSA* **73**(3), 332–338. Cited on page 193.

Khan, M., Trujano, F., Choudhury, A. and Maes, P. (2018), Mathland: Playful mathematical learning in mixed reality, *in* 'Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems', ACM, p. D108. Cited on page 175.

Khoshelham, K. and Elberink, S. O. (2012), 'Accuracy and resolution of kinect depth data for indoor mapping applications', *Sensors* **12**(2), 1437–1454. Cited on page 58.

Kim, J.-S. and Park, J.-M. (2015), Physics-based hand interaction with virtual objects, *in* 'Robotics and Automation (ICRA), 2015 IEEE International Conference on', IEEE, pp. 3814–3819. Cited on pages XV, 192, and 193.

Kimura, Z. and Sato, M. (2018), An examination on effective auditory stimulation when grasping a virtual object with a bare hand, *in* 'Advanced Image Technology (IWAIT), 2018 International Workshop on', IEEE, pp. 1–4. Cited on pages 192, 208, and 210.

Kitsikidis, A., Dimitropoulos, K., Douka, S. and Grammalidis, N. (2014), Dance analysis using multiple kinect sensors, *in* 'Computer Vision Theory and Applications (VISAPP), 2014 International Conference on', Vol. 2, IEEE, pp. 789–795. Cited on page 25.

Klatzky, R. L., Lederman, S. J. and Reed, C. (1987), 'There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision.', *Journal of experimental psychology: general* **116**(4), 356. Cited on page 58.

Klein, A. and De Assis, G. A. (2013), A markeless augmented reality tracking for enhancing the user interaction during virtual rehabilitation, *in* 'Virtual and Augmented Reality (SVR), 2013 XV Symposium on', IEEE, pp. 117–124. Cited on page 23.

Kruijff, E., Swan, J. E. and Feiner, S. (2010), Perceptual issues in augmented reality revisited, *in* 'Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on', IEEE, pp. 3–12. Cited on page 193.

Kruskal, W. H. and Wallis, W. A. (1952), 'Use of ranks in one-criterion variance analysis', *Journal of the American statistical Association* **47**(260), 583–621. Cited on pages 68 and 103.

Landsmeer, J. (1962), 'Power grip and precision handling', *Annals of the rheumatic diseases* **21**(2), 164. Cited on pages X, 35, 36, and 201.

Lee, G. A .and Billinghurst, M. and Kim, G. J. (2004), 'Occlusion based interaction methods for tangible augmented reality environments', *Virtual Reality Continuum And Its Applications* **1**(212), 419. Cited on pages IX, 17, and 18.

Lee, M., Green, R. and Billinghurst, M. (2008), 3d natural hand interaction for ar applications, *in* 'Image and Vision Computing New Zealand, 2008. IVCNZ 2008. 23rd International Conference', IEEE, pp. 1–6. Cited on page 25.

León, B., Morales, A. and Sancho-Bru, J. (2014), *From robot to human grasping simulation*, Springer. Cited on pages X, 30, 31, 42, and 44.

Levin, M. F., Magdalon, E. C., Michaelsen, S. M. and Quevedo, A. A. (2015), 'Quality of grasping and the role of haptics in a 3-d immersive virtual reality environment in individuals with stroke', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **23**(6), 1047–1055. Cited on page 15.

Lindeman, R. W., Lee, G., Beattie, L., Gamper, H., Pathinarupothi, R. and Akhilesh, A. (2012), Geoboids: A mobile ar application for exergaming, *in* 'Mixed and Augmented Reality (ISMAR-AMH), 2012 IEEE International Symposium on', IEEE, pp. 93–94. Cited on page 14.

Liu, H., Ju, Z., Ji, X., Chan, C. S. and Khoury, M. (2017), Human hand motion analysis with multisensory information, *in* 'Human Motion Sensing and Recognition', Springer, pp. 171–191. Cited on page 41.

Liu, P.-P. and Wu, Y.-C. (2018), Research on real-time graphics drawings technology in virtual scene, *in* '2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)', IEEE, pp. 1–6. Cited on page 57.

Liu, S. L. E. and Cheung, H. K. C. (2016), 'Development of practical vocational training class making use of virtual reality-based simulation system and augmented reality technologies'. Cited on page 208.

Lun, R. and Zhao, W. (2018), Kinect applications in healthcare, *in* 'Encyclopedia of Information Science and Technology, Fourth Edition', IGI Global, pp. 5876–5885. Cited on page 58.

Luo, X., Kline, T., Fischer, H. C., Stubblefield, K. A., Kenyon, R. V. and Kamper, D. G. (2006), Integration of augmented reality and assistive devices for post-stroke hand opening rehabilitation, *in* '2005 IEEE Engineering in Medicine and Biology 27th Annual Conference', IEEE, pp. 6855–6858. Cited on page 188.

MacKenzie, C. L. and Iberall, T. (1994), *The grasping hand*, Vol. 104, Elsevier. Cited on pages X, XI, 14, 30, 31, 32, 33, 34, 35, 36, 38, 40, 41, 42, 44, 50, 58, 188, 201, and 206.

MacKenzie, C. L. and Marteniuk, R. G. (1985), 'Motor skill: Feedback, knowledge, and structural issues.', *Canadian Journal of Psychology/Revue canadienne de psychologie* **39**(2), 313. Cited on page 42.

Magdalon, E. C., Michaelsen, S. M., Quevedo, A. A. and Levin, M. F. (2011), 'Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment', *Acta psychologica* **138**(1), 126–134. Cited on pages 15 and 189.

Maisto, M., Pacchierotti, C., Chinello, F., Salvietti, G., De Luca, A. and Prattichizzo, D. (2017), 'Evaluation of wearable haptic systems for the fingers in augmented reality applications', *IEEE Transactions on Haptics* . Cited on page 11.

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* pp. 50–60. Cited on page 179.

Maria, K., Filippeschi, A., Ruffaldi, E., Shorr, Y. and Gopher, D. (2015), Evaluation of multimodal feedback effects on the time-course of motor learning in multimodal vr platform for rowing training, *in* 'IEEE ICVR', pp. 158–159. Cited on page 98.

Marteniuk, R., MacKenzie, C., Jeannerod, M., Athenes, S. and Dugas, C. (1987), 'Constraints on human arm movement trajectories.', *Canadian Journal of Psychology/Revue canadienne de psychologie* **41**(3), 365. Cited on page 42.

Mei, J., Liu, J., Zhang, X., Lu, X. and Huang, J. (2017), 'Augmented reality-based training system for hand rehabilitation', *Multimedia Tools and Applications* **76**(13), 14847–14867. Cited on page 188.

Méndez, R., Flores, J., Castelló, E. and Arenas, R. (2016), Preliminary evaluation of the kinect v2 sensor for its use in virtual tv sets with natural interaction, *in* 'Proceedings of the XVII International Conference on Human Computer Interaction', ACM, p. 14. Cited on pages 10 and 28.

Milgram, P. and Colquhoun, H. (1999), 'A taxonomy of real and virtual world display integration', *Mixed reality: Merging real and virtual worlds* **1**, 1–26. Cited on pages IX and 9.

Milgram, P. and Kishino, F. (1994), 'A taxonomy of mixed reality visual displays', *IEICE TRANSACTIONS on Information and Systems* **77**(12), 1321–1329. Cited on pages 9 and 10.

Milgram, P., Takemura, H., Utsumi, A. and Kishino, F. (1995), Augmented reality: A class of displays on the reality-virtuality continuum, *in* 'Telemanipulator and telepresence technologies', Vol. 2351, International Society for Optics and Photonics, pp. 282–293. Cited on page 10.

Moehring, M. and Froehlich, B. (2011), Effective manipulation of virtual objects within arm's reach, *in* 'Virtual Reality Conference (VR), 2011 IEEE', IEEE, pp. 131–138. Cited on page 16.

Moeslund, T. B., Hilton, A. and Krüger, V. (2006), 'A survey of advances in vision-based human motion capture and analysis', *Computer vision and image understanding* **104**(2), 90–126. Cited on page 24.

Mon-Williams, M. and Tresilian, J. R. (2000), 'Ordinal depth information from accommodation?', *Ergonomics* **43**(3), 391–404. Cited on page 194.

Napier, J. R. (1956), 'The prehensile movements of the human hand', *Bone & Joint Journal* **38**(4), 902–913. Cited on pages X, 35, 36, 37, 43, and 200.

Newell, K., Scully, D., Tenenbaum, F. and Hardiman, S. (1989), 'Body scale and the development of prehension', *Developmental psychobiology* **22**(1), 1–13. Cited on page 59.

Nicolau, S., Soler, L., Mutter, D. and Marescaux, J. (2011), 'Augmented reality in laparoscopic surgical oncology', *Surgical oncology* **20**(3), 189–201. Cited on page 2.

Nowak, D. A. and Hermsdörfer, J. (2009), *Sensorimotor control of grasping: physiology and pathophysiology*, Cambridge University Press. Cited on pages 30, 31, 33, and 34.

Oldfield, R. C. (1971), 'The assessment and analysis of handedness: the edinburgh inventory', *Neuropsychologia* **9**(1), 97–113. Cited on page 62.

Oliveira, E., Simões, F. P. and Correia, W. F. (2017), Heuristics evaluation and improvements for low-cost virtual reality, *in* '2017 19th Symposium on Virtual and Augmented Reality (SVR)', IEEE, pp. 178–187. Cited on page 23.

Pacchierotti, C., Chinello, F., Malvezzi, M., Meli, L. and Prattichizzo, D. (2012), Two finger grasping simulation with cutaneous and kinesthetic force feedback, *in* 'Haptics: perception, devices, mobility, and communication', Springer, pp. 373–382. Cited on page 2.

Pande, B. and Singh, I. (1971), 'One-sided dominance in the upper limbs of human fetuses as evidenced by asymmetry in muscle and bone weight.', *Journal of Anatomy* **109**(Pt 3), 457. Cited on page 62.

Parry, C. W. (1966), *Rehabilitation of the Hand*, Butterworth. Cited on page 36.

Patkin, M. (1981), 'Ergonomics in microsurgery', *Australian and New Zealand Journal of Obstetrics and Gynaecology* **21**(3), 134–136. Cited on page 36.

Pham, T.-H., Kyriazis, N., Argyros, A. A. and Kheddar, A. (2018), 'Hand-object contact force estimation from markerless visual tracking', *IEEE transactions on pattern analysis and machine intelligence* **40**(12), 2883–2896. Cited on pages 23 and 57.

Pitts, M. J., Burnett, G., Skrypchuk, L., Wellings, T., Attridge, A. and Williams, M. A. (2012), 'Visual–haptic feedback interaction in automotive touchscreens', *Displays* **33**(1), 7–16. Cited on page 98.

Piumsomboon, T., Clark, A., Billinghurst, M. and Cockburn, A. (2013), User-defined gestures for augmented reality, *in* 'CHI'13 Extended Abstracts on Human Factors in Computing Systems', ACM, pp. 955–960. Cited on pages 49, 198, 199, and 200.

Ponto, K., Kimmel, R., Kohlmann, J., Bartholomew, A. and Radwir, R. G. (2012), 'Virtual exertions: A user interface combining visual information, kinesthetics and biofeedback for virtual object manipulation', *Proceedings - IEEE Symposium on 3D User Interfaces* pp. 85–88. Cited on pages 8, 14, and 23.

Prattichizzo, D., Pacchierotti, C. and Rosati, G. (2012), 'Cutaneous force feedback as a sensory subtraction technique in haptics', *Haptics, IEEE Transactions on* **5**(4), 289–300. Cited on pages 60 and 99.

Rautaray, S. S. and Agrawal, A. (2015), 'Vision based hand gesture recognition for human computer interaction: a survey', *Artificial Intelligence Review* **43**(1), 1–54. Cited on page 23.

Reither, L. R., Foreman, M. H., Migotsky, N., Haddix, C. and Engsberg, J. R. (2018), 'Upper extremity movement reliability and validity of the kinect version 2', *Disability and Rehabilitation: Assistive Technology* **13**(1), 54–59. Cited on page 58.

Rekimoto, J. (2002), 'Smartskin: an infrastructure for freehand manipulation on interactive surfaces', *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* **02**, 113–120. Cited on pages IX, 17, and 18.

Ren, D., Goldschwendt, T., Chang, Y. and Höllerer, T. (2016), Evaluating wide-field-of-view augmented reality with mixed reality simulation, *in* 'Virtual Reality (VR), 2016 IEEE', IEEE, pp. 93–102. Cited on page 172.

Rijpkema, H. and Girard, M. (1991), Computer animation of knowledge-based human grasping, *in* 'ACM Siggraph Computer Graphics', Vol. 25, ACM, pp. 339–348. Cited on page 44.

Ro, J. Y., Debowy, D., Ghosh, S. and Gardner, E. P. (2000), 'Depression of neuronal firing rates in somatosensory and posterior parietal cortex during object acquisition in a prehension task', *Experimental Brain Research* **135**(1), 1–11. Cited on page 44.

Rogers, Y., Sharp, H. and Preece, J. (2011), *Interaction design: beyond human-computer interaction*, John Wiley & Sons. Cited on pages 9, 12, and 23.

Roth, I. and Frisby, J. P. (1986), *Perception and representation: a cognitive approach*, Open University. Cited on page 9.

Sattler, M., Sarlette, R., Mücken, T. and Klein, R. (2005), Exploitation of human shadow perception for fast shadow rendering, *in* 'Proceedings of the 2nd symposium on Applied perception in graphics and visualization', ACM, pp. 131–134. Cited on page 139.

Schlesinger, I. G. (1919), 'Der mechanische aufbau der kunstlischen glieder [the mechanical structure of artificial limbs]', *M. Borchardt et al. (Eds.), Ersatznlieder und Arbeitshilfen fur Kriensbeschadiete und Unfallverletzte* pp. 21–600. Cited on pages X, 34, and 35.

Schwarz, R. J. and Taylor, C. (1955), 'The anatomy and mechanics of the human hand', *Artificial limbs* **2**(2), 22–35. Cited on pages X and 35.

Slocum, D. B. and Pratt, D. R. (1946), 'Disability evaluation for the hand.', *JBJS* **28**(3), 491–495. Cited on page 35.

Sollerman, C. (1980), 'Grip function of the hand. analysis evaluation and a new test method', *Section of Hand Surgery, Department of Orthopaedic Surgery, Sahlgren Hospital, University of Goteborg, Goteberg, Sweden* . Cited on page 40.

Stockmeier, K., Horton, H. and Franz, V. H. (2003), 'How do we grasp (virtual) objects in three-dimensional space?', *Journal of Vision* **3**(9), 383–383. Cited on page 56.

Sturman, D. J. and Zeltzer, D. (1994), 'A survey of glove-based input', *IEEE Computer graphics and Applications* **14**(1), 30–39. Cited on page 27.

Supuk, T., Bajd, T. and Kurillo, G. (2011), 'Assessment of reach-to-grasp trajectories toward stationary objects', *Clinical biomechanics* **26**(8), 811–818. Cited on page 30.

Sutherland, C., Hashtrudi-Zaad, K., Sellens, R., Abolmaesumi, P. and Mousavi, P. (2013), 'An augmented reality haptic training simulator for spinal needle procedures', *Biomedical Engineering, IEEE Transactions on* **60**(11), 3009–3018. Cited on page 2.

Suzuki, S., Suzuki, H. and Sato, M. (2014), 'Grasping a virtual object with a bare hand', *ACM SIGGRAPH 2014 Posters on - SIGGRAPH '14* pp. 1–1. Cited on pages 8, 21, 22, and 23.

Swan, J. E., Kuparinen, L., Rapson, S. and Sandor, C. (2017), 'Visually perceived distance judgments: Tablet-based augmented reality versus the real world', *International Journal of Human–Computer Interaction* **33**(7), 576–591. Cited on pages 190, 193, and 197.

Swan, J. E., Singh, G. and Ellis, S. R. (2015), 'Matching and Reaching Depth Judgments with Real and Augmented Reality Targets', *IEEE Transactions on Visualization and Computer Graphics* **21**(11), 1289–1298. Cited on pages XV, 66, 192, 193, 195, and 197.

Terpstra, T. J. (1952), 'The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking', *Indagationes Mathematicae* **14**(3), 327–333. Cited on page 68.

Tsoupikova, D., Stoykov, N. S., Corrigan, M., Thielbar, K., Vick, R., Li, Y., Triandafilou, K., Preuss, F. and Kamper, D. (2015), 'Virtual immersion for post-stroke hand rehabilitation therapy', *Annals of biomedical engineering* **43**(2), 467–477. Cited on pages 15 and 16.

Uno, Y., Fukumura, N., Suzuki, R. and Kawato, M. (1993), Integration of visual and somatosensory information for preshaping hand in grasping movements, *in* 'Advances in Neural Information Processing Systems', pp. 311–318. Cited on page 40.

Van Krevelen, D. and Poelman, R. (2010), 'A survey of augmented reality technologies, applications and limitations', *International Journal of Virtual Reality* **9**(2), 1. Cited on pages 10 and 12.

Vieira, J., Sousa, M., Arsénio, A. and Jorge, J. (2015), Augmented reality for rehabilitation using multimodal feedback, *in* 'Proceedings - Workshop on ICTs for improving Patients Rehabilitation Research Techniques', ACM, pp. 38–41. Cited on pages 2 and 99.

Wade, N. J. and Swanston, M. (2013), *Visual perception: An introduction*, Psychology Press. Cited on pages 61 and 193.

Wilcoxon, F. and Wilcox, R. A. (1964), *Some rapid approximate statistical procedures*, Lederle Laboratories. Cited on page 142.

Winters, J. M. and Crago, P. E. (2012), *Biomechanics and neural control of posture and movement*, Springer Science & Business Media. Cited on pages 29, 30, 32, and 33.

Wobbrock, J. O., Morris, M. R. and Wilson, A. D. (2009), User-defined gestures for surface computing, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', ACM, pp. 1083–1092. Cited on page 49.

Woodworth, R. S. (1899), *The accuracy of voluntary movement...*, Columbia University. Cited on page 43.

Wu, M. and Balakrishnan, R. (2003), 'Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays', *Proceedings of the 16th annual ACM symposium on User interface software and technology* **5**(2), pages 193–202. Cited on pages IX, 17, and 18.

Yang, L., Zhang, L., Dong, H., Alelaiwi, A. and El Saddik, A. (2015), 'Evaluating and improving the depth accuracy of kinect for windows v2', *IEEE Sensors Journal* **15**(8), 4275–4285. Cited on page 58.

Yang, X., Yeh, S.-C., Niu, J., Gong, Y. and Yang, G. (2017), Hand rehabilitation using virtual reality electromyography signals, *in* '2017 5th International Conference on Enterprise Systems (ES)', IEEE, pp. 125–131. Cited on page 14.

Zhou, F., Duh, H. B.-L. and Billinghurst, M. (2008), Trends in augmented reality tracking, interaction and display: A review of ten years of ismar, *in* 'Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality', IEEE Computer Society, pp. 193–202. Cited on pages 12 and 26.