

106308

2
2006

**LIBRARY
Michigan State
University**

This is to certify that the
dissertation entitled

**SPATIAL INFORMATION DISPLAY FRAMEWORK FOR
MOBILE AUGMENTED REALITY INTERFACE**

presented by

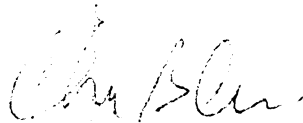
Kwok Hung Tang

has been accepted towards fulfillment
of the requirements for the

Ph.D.

degree in

**Computer Science and
Engineering**



Major Professor's Signature

1st November 2005

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
JUN 21 2007		

**SPATIAL INFORMATION DISPLAY FRAMEWORK FOR MOBILE
AUGMENTED REALITY INTERFACES**

By

Kwok Hung Tang

**A DISSERTATION
Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of**

DOCTOR OF PHILOSOPHY

Department of Computer Science and Engineering

2005

ABSTRACT

SPATIAL INFORMATION DISPLAY FRAMEWORK FOR MOBILE AUGMENTED REALITY INTERFACES

By

Kwok Hung Tang

Future augmented reality (AR) user interfaces will allow designers the flexibility of placing information all around the body of a mobile user, effectively utilizing the area around the body as a spatial user interface. The design of these future interfaces prompts a significant human factors challenge: *How should interface designers map different metaphors, information, and functions of computer usage into a volumetric computing environment to maximize information bandwidth and reduce a user's attentional and cognitive load?* Issues of human cognition and psychological effects in AR are mainly unexplored, and little is known about how humans organize information objects in an egocentric and exocentric free-space environment. This thesis addresses the research problem by: (1) constructing a spatial information display framework based on neuropsychological research, and (2) extending research in cognitive psychology and behavioral science to AR interface design. Three research questions in cognitive psychology are identified that are closely related to the design of AR interfaces: (1) the use of reference frames during the spatial encoding process, (2) applicability of perceptual asymmetry properties in AR interface design, and (3) directing visuo-spatial attention in omnidirectional

space. Six experiments were conducted to investigate these three research questions. The experimental results were combined with existing literature to form a set of information display guidelines for information display in mobile AR environments.

Keywords:

Augmented reality, human-computer interaction, perceptual and kinematics asymmetry, spatial reference frame, three-dimensional visuo-spatial attention

ACKNOWLEDGEMENTS

The thesis is the collaborative effort, and would not have materialized without the help of my advisors, colleagues, friends and family. I own more to my parents than anyone else. It is to them my thesis is dedicated to. I am deeply indebted to my two academic fathers, Dr. Charles Owen and Dr. Frank Biocca. It was a great pleasure to be a student of Dr. Charles Owen, my principle advisor. His excellent consultancy, professional editing, unrelenting support, and hard working attitude not only provide a solid basis for my thesis research, but also set a great role model for my future career. It has been a prestigious opportunity to be able to work with Dr. Frank Biocca in the M.I.N.D. Labs for the last seven years. His philosophical mind has been of great value to the theoretical background of my thesis research, and his vision has broadened and deepened my view on scientific research. My sincere gratitude goes to Dr. George Stockman and Dr. John Weng for monitoring my research work and spending their valuable time in reading the document and providing valuable advisor to this thesis. I wish to express my warm and sincere thanks to Dr. Weimin Mou for his contribution of a significant amount of research work in this thesis. My sincere thanks also go to Dr. Prabu David for his statistical analysis on Experiment 5. I also owe my fellow colleague, Fan Xiao, a big thank for his support in preparing the stimulus materials. I must also thank Betsy McKeon, who did a very competent work on data collection in the experiments. And last but not least, I would like to express a warm gratitude to Zena Biocca, who created a great working environment in the lab to foster this thesis research.

TABLE OF CONTENTS

LIST OF TABLES.....	viii
LIST OF ABBREVIATION.....	xii
1 Introduction.....	1
1.1 Using Space as a Medium for Thought.....	2
1.2 How Spatial Representations Leverage Spatial Cognition for Thinking.....	3
1.3 Spatial Cognition and Augmented Reality Space.....	5
1.4 Research Motivation and Problem Statement.....	6
1.5 Contributions of this Thesis.....	7
2 Theoretical Background.....	9
2.1 Spatial Framework of Three-dimensional Space.....	9
2.2 Neuropsychology of Three-dimensional Spaces	10
2.2.1 Personal/Body Space	10
2.2.2 Peripersonal Space.....	12
2.2.3 Extrapersonal space	14
2.3 Mapping Digital Information to Space in Augmented Reality Systems.....	17
3 Spatial Framework of Information Display in Mobile Augmented Reality Environments	19
3.1 Spatial Information Framework.....	19
3.1.1 Personal-body Infospace.....	19
3.1.2 Peripersonal Infospace.....	23
3.1.3 Extrapersonal Focal Infospace.....	25
3.1.4 Extrapersonal Action-Scene Infospaces	25
3.1.5 Extrapersonal Ambient Infospaces	25
3.2 Summary.....	26
4 Behavioral Properties of Three-dimensional Space.....	28
4.1 Behavioral Properties in Personal-body Infospaces	28
4.1.1 Proprioception.....	29
4.1.2 Spatial Bias in Personal-body Infospace.....	31
4.1.3 The Hand and Forearms Personal-body Infospaces.....	31
4.2 Behavioral Properties in Peripersonal Infospace	33
4.2.1 Spatial Biases of Information in Peripersonal Space.....	33
4.3 Behavioral Properties in Extrapersonal Focal Infospace	34
4.3.1 Visual Clutter in Head Stabilized Reference Frame.....	35
4.3.2 Perceptual Fading of Visual Stimulus.....	35
4.3.3 Spatial Bias in Head Stabilized Reference Frame	36
4.4 Behavioral Properties in Relation to Egocentric Infospaces.....	36
4.4.1 Kinematics Asymmetry	37
4.4.2 Perceptual Asymmetries	39

4.5	Behavioral Properties in Extrapersonal Action-scene Infospaces	45
4.5.1	Spatial Consistency of Information Objects with the Environment	46
4.5.2	Remote Interaction for Information Objects in Extrapersonal Action-scene Infospace: Selection and Manipulation	47
4.5.3	Unregistered Extrapersonal Action-scene Infospace	53
4.6	Behavioral Properties in Extrapersonal Ambient Infospace	53
4.6.1	Spatial Bias in Extrapersonal Ambient Infospace	53
4.6.2	Linear Perspective and Motion Perception Properties	54
4.7	Summary	54
5	Reference Frames in Mobile Augmented Reality Displays	55
5.1	Related Works	55
5.2	Experiment 1: The Default Reference Frame	59
5.2.1	Methodology	59
5.3	Experiment 2: Adaptation of Egocentric Frame with Prior Experience	66
5.3.1	Methodology	67
5.3.2	Results and Discussion	68
5.4	Experiment 3: Adaptation to an Egocentric Frame with Oral Instruction	70
5.4.1	Methodology	70
5.5	Discussion	72
6	Evaluation of Perceptual Asymmetric Effects in Egocentric Infospaces	76
6.1	Experiment 4: Evaluation of Left vs. Right Instruction Presentation	76
6.1.1	Methodology	77
6.1.2	Results	79
6.1.3	Discussion	80
6.2	Experiment 6: Emotion and Semantic Meaning	81
6.2.1	Related Works	81
6.2.2	Methodology	83
6.2.3	Results and Analysis	86
6.2.4	Discussion	92
6.3	Summary	95
7	Directing Attention in Mobile AR Interface	96
7.1	Attention Management	98
7.1.1	Attention Cueing in Existing Interfaces	99
7.1.2	Spatial Cueing in Augmented Reality	100
7.2	The Omnidirectional Attention Funnel	101
7.2.1	Components of the Attention Funnel	102
7.2.2	Affordances in the Attention Funnel that Guide Navigation and Body Rotation	106
7.2.3	Methods for Sensing or Marking Targets Objects or Locations	107
7.3	Methodology	108
7.3.1	Participants	109
7.3.2	Stimulus Materials	109
7.3.3	Apparatus and Test Environment	110

7.3.4	Measurements	111
7.3.5	Procedure	111
7.4	Results.....	112
7.5	Discussion.....	114
7.6	Application of the Attention Funnel	114
8	Discussion and Conclusion.....	117
8.1	Guideline for Information Display in Augmented Reality Environments.....	118
8.2	Future Works	119
8.3	Conclusion	120
Appendix A. Spatial Information Display Guideline for Mobile Augmented Reality		
	Interfaces.....	122
A.	Spatial Framework of the Three-dimensional Space	123
B.	Peripersonal Infospace	124
C.	Personal-body Infospace.....	126
D.	Extrapersonal Focal Infospace.....	129
E.	Egocentric Infospaces	132
F.	Extrapersonal Action-scene Infospace.....	135
G.	Extrapersonal Ambient Infospace.....	138
H.	Infospace Choice for Common Information Objects.....	139
9	References.....	142

LIST OF TABLES

Table 4.1. Summary of cerebral hemispheric specializations.....	41
Table 4.2. Summary of pros and cons of different remote objects manipulation method	52
Table 5.1. Pointing latency (in seconds) and pointing accuracy (in degrees) as a function of Actual-Imagined (A-I) distance and Learning-Imagined (L-I) distance in Experiment 1.....	64
Table 5.2. Analysis of variance results for pointing latency and pointing accuracy in Actual-Imagined (A-I) and Learning-Imagined (L-I) conditions in Experiment 1.....	65
Table 5.3. Pointing latency (in seconds) and pointing accuracy (in degrees) as a function of Actual-Imagined (A-I) distance and Learning-Imagined (L-I) distance in Experiment 2.....	68
Table 5.4. Analysis of variance results for pointing latency and pointing accuracy in Actual-Imagined (A-I) and Learning-Imagined (L-I) conditions in Experiment 2.....	69
Table 5.5. Pointing latency (in seconds) and pointing accuracy (in degrees) as a function of Actual-Imagined (A-I) distance and Learning-Imagined (L-I) distance in Experiment 3.....	71
Table 5.6. Analysis of variance results for pointing latency and pointing accuracy in Actual-Imagined (A-I) and Learning-Imagined (L-I) conditions in Experiment 3.....	72
Table 6.1. Task completion time and standard deviation in Experiment 4.....	80
Table 6.2. Means for the Different levels of the 3 Experimental Factors.....	87

LIST OF FIGURES

Figure 2.1.	A prototype volumetric AR interface with information objects placed in different reference frame.	10
Figure 2.2	Personal/body space.....	11
Figure 2.3.	Peripersonal space.....	13
Figure 2.4.	Extrapersonal focal space.	14
Figure 2.5.	Extrapersonal action space.....	15
Figure 2.6.	Scene space.	16
Figure 2.7.	Extrapersonal ambient space.....	17
Figure 3.1.	Human skeletal structure (Gray, Bannister, Berry and Williams 1995).	21
Figure 3.2.	Selected bone groups for Personal-Body Infospaces: (a) Skull (b) Vertebral Column (c) Sternum and costal cartilages (d) Humerus (e) Forearm group (f) Hand (g) Femur (h) Patella (i) Leg (j) Foot (Gray et al. 1995).....	22
Figure 3.3.	Skeletal structure of human hand (Gray et al. 1995).	23
Figure 3.4.	The vertebral column (Gray et al. 1995).....	24
Figure 3.5.	Spatial framework for information display in mobile augmented reality environments.....	27
Figure 4.1.	Hand Personal-body Infospaces: a menu attached to the non-dominant hand and an interaction tool (the ring) attached to the dominant hand.	32
Figure 4.2.	The visual pathway of left and right visual field. Retinal signal from the left and right visual fields projects exclusively to the contralateral cerebral hemispheres.	42
Figure 4.3.	Visual pathway of the upper and lower visual field.	44
Figure 5.1.	The eight virtual objects used in the experiments.....	60
Figure 5.2.	Layout of objects used in the experiments. During the learning phase, half of the participants faced the cell phone and the other half faced the notebook.	60
Figure 5.3.	Design of experiments: Head-nose icons indicate actual headings; arrows indicate imagined headings. Headings and differences between them are	

	measured counter-clockwise to maintain consistency with previous experiments.....	61
Figure 6.1.	Examples of instruction and the completed task. Example of text instruction is shown in (a) instruction and the completed task is shown in (b), and example of graphic instruction is shown in (c) and the completed task is shown in (d).	78
Figure 6.2.	Ten predefined locations around the body. The five locations in the near space are 3' away from the body. The five locations in the far space is 10' from the body. The above, below, left and right locations is deviated 30° from the center location.	84
Figure 6.3.	The two stimulus material used in the experiment. The golden sphere used for object representation is shown on the left side. The human head used for agent representation is shown on the right side.	84
Figure 6.4.	Relevant-Irrelevant by distance and position of objects	89
Figure 6.5.	Urgent-Not urgent by distance and position.	91
Figure 6.6.	Urgent-Not Urgent by object and position.....	91
Figure 7.1.	The attention funnel links the head of the viewer directly to an object anywhere around the body.....	101
Figure 7.2.	Three basic patterns are used to construct a funnel: (A) the head centered plane includes a boresight to mark the center of the pattern from the user's viewpoint, (B) funnel planes, added in a fixed pattern (approximately every 12 centimeters) between the user and the object, and (C) the object marker pattern that includes a red cross hairs marking the approximate center of the object.....	103
Figure 7.3.	As the head and body move, the attention funnel dynamically provides continuous feedback. Affordances from the perspective cues automatically guide the user towards the cued location or object. Dynamic head movement cues are provided by the skew (e.g., left, right, up, down) of the attention funnel. The level of alignment (skew) of the funnel provides an immediate intuitive sense of how much the body or head must turn to see the object.	104
Figure 7.4.	Example of the attentional funnel drawing attention of the user to an object on the shelf, the red box.....	107
Figure 7.5.	Test Environment: The user sat in the middle of test environment for the visual search task. It consisted of an omnidirectional workspace assembled from four tables each with 12 objects (6 primitive shapes and 6 general office objects) for a total of 48 target search objects.....	110

Figure 7.6	Search time and consistency by experimental condition. Attentional funnel decreased search time by 22% on average (28% when reach time is subtracted) and increased search consistency (decreased variability) by 65%.	113
Figure 7.7.	Mental workload measured by NASA TLX for each experimental condition.	113

LIST OF ABBREVIATION

A-I	Actual-Imagined
AR	Augmented Reality
ANOVA	Analysis of Variance
B.C.E.	Before the Common Error
CHIMP	Chapel Hill Immersive Modeling Program
GPS	Global Positioning System
HMD	Head-mounted Display
<i>HOMER</i>	Hand-centered Object Manipulation Extending Ray-casting
HUD	Head-up Display
LGN	Lateral Geniculate Nucleus
L-I	Learning-Imagined
M	Mean
ms	Milliseconds
PDA	Personal Digital Assistant
RFID	Radio Frequency Identification
SD	Standard Deviation
SGI	Silicon Graphics, Inc.
VR	Virtual Reality
WIMP	Window, Icon, Menu, Pointer

1 Introduction

Technological developments are allowing for the design of computer user interfaces that extend the traditional interface into the physical space all around the user, breaking the bounds of the small monitor-based display and allowing for mobile interfaces that appear to present a virtually unlimited quantity of information objects around the user. User interface components can float in space around the user or appear to be placed on the surface of the body. This extension of the space utilized for information brings to question how best to place content around the user. This thesis explores the effective utilization of the space around a user in future user interfaces, addressing issues of effective placement that are sound from a psychological and physiological standpoint.

Alan Kay described the personal computer as the first meta-medium – an electronic medium which can be used to store, manipulate and access numerous media forms such as text, images, audio, video, and three-dimensional models (Kay 1984). The emergence of the World Wide Web in the last decade brought into existence the “global village interconnected by an electronic nervous system” as envisioned by Marshall McLuhan (1967). During this era, the computer has evolved into an information portal to databases in different media forms and a communication portal for different social activities. An unprecedented amount of information and activities can be received continuously through this portal by the user. The *user interface* is analogous to a gateway for this communication and information portal. It manages, and often limits, the information the user is able to absorb and the commands the user is able to deploy to the

computer system. Effective design of this gateway can maximize the bandwidth between the computer and the user.

1.1 Using Space as a Medium for Thought

Every medium, from traditional printed media to modern computer-mediated interactive media, uses spatial arrangement in some way to organize information (Cavell 2002). The prevalent computer user interface for the last 25 years, the traditional WIMP (window, icon, menu, pointer) direct manipulation interface (Shneiderman 1983) is a two-dimensional spatial arrangement of icons and overlapping windows suggesting layers of information and containers (or folders) that are “opened” to reveal arrays of icons and simulating the arrangement of material as if it were on an office desktop. Motor interaction in WIMP interfaces is spatial, as the system is controlled by a virtual pointer on the display manipulated by the mouse on a spatial surface. The advantage of the WIMP interface is familiarity. It is based on the desk surface and folders metaphor that is obvious to novice users. However the metaphor is limited in much the same way limiting an office to just the surface of a small desk would be. Three-dimensional environments are far richer and more expressive than two-dimensional flat surfaces.

With the advent of motion tracking systems and low-cost, high-performance graphics workstations, the novel and highly spatial augmented reality (AR) interfaces visualized in Hollywood movies, video games and science fiction are becoming technologically feasible. These interfaces tightly couple spatial three-dimensional stimuli to the movement of the user’s body. The sensors and effectors of the computer system are then mapped to the user’s body schema (Biocca 1997). Volumetric AR interfaces make use of a greater range of human sensorimotor capabilities, potentially increasing the

communication bandwidth between the user and the computer by cutting the ties to that technological ancestor – the typewriter.

AR interfaces have very unique characteristics as compared to other media and computer interfaces: users interact with the computer system through body motion in a volumetric space, instead of via a two-dimensional surface. This is very different from traditional computer interfaces and other three dimensional screen-based interactions such as DataMountain (Robertson, Czerwinski, Larson, Robbins, Thiel and van Dantzich 1998) and fish-tank virtual reality (VR) (Ware, Arthur and Booth 1993). Traditional computer interfaces can be likened to limiting user interaction to the surface area of a small office desktop, and screen-based three-dimensional interfaces are analogous to a window into the office through which users peer at a presentation of an alternative reality; effectively an outsider looking in. AR is a truly immersive spatial electronic medium in which the user's body is immersed into a blended real/virtual environment, where the computer arrays two-dimensional and three-dimensional information around the user. This unique spatial arrangement allows for the display of large volumes of data, and designers are still exploring ways to organize information in this cutting edge interface.

1.2 How Spatial Representations Leverage Spatial Cognition for Thinking

In the everyday world, humans organize and manipulate objects in space to facilitate thinking. Kirsh asserted that humans are constantly, whether consciously or subconsciously, organizing and reorganizing space in everyday life to enhance performance, and argued that “methods used to manage our space are key to organization of our thought patterns and behavior” (Kirsh 1995).

Spatial schema and spatial reasoning are not just about space. They are also implicated in abstract reasoning. There is ample evidence from the fields of psychology and neuroscience that spatial cognition plays an important role in mathematical reasoning, modeling of time, language organization, and memory organization (Gardner 1983; Bryant 1992; Bryant, Tversky and Franklin 1992; Kirsh and Maglio 1992; Eilan, McCarthy and Brewer 1993; Ferguson 1994; Grabowska and Nowicka 1996; Boroditsky and Ramscar 2002). The use of spatial representation and organization to enhance human cognition has been a successful strategy since the effective mnemonic strategies of the ancient Greeks. Demosthenes, a Greek orator born around 384 B.C.E., used a strategy known as “Method of Loci” to memorize long speeches by mentally walking through his house, associating each element in the speech with different spots or objects in the house (Yates 1966).

How information is spatially represented can facilitate cognition. For example, different spatial arrangement of physical objects can dramatically affect how people solve a problem. Zhang and Norman reported an experiment showing that a subject's performance when solving the Tower of Hanoi problem was drastically affected by the spatial placement of the problem pieces (Zhang and Norman 1994). Much of the problem representation of the Tower of Hanoi problem can be offloaded to an external spatial representation of the problem pieces, and as a result, the load on internal working memory can be reduced and more working memory capacity can be allocated to problem solving.

There is historical evidence that the arrival of new ways to visualize information, such as illustrations, graphs, computer graphics and videos, has had a dramatic impact on

advances in engineering and science (Ferguson 1994). Virtual environments and visualizations represent information spatially through proximity, color gradation, or spatial arrays to allow users to immediately grasp large amounts of quantitative data and complex mathematical relationships (Card, Mackinlay and Shneiderman 1999; Ware 2000). Spatial arrays can be intuitive for even novel users. For example, Merickel found that VR enhanced a child's ability to solve spatially related problems (Merickel 1992).

1.3 Spatial Cognition and Augmented Reality Space

Wearable and mobile AR systems have a great potential to provide continuous support for virtual space and visualized information arrays, as well as integrating, annotating, and interacting with physical space. These systems can potentially be powerful “cognitive artifacts” (Norman 1993) or “intelligence amplifying systems” (Brooks 1996) that enhance human cognitive activities, such as attention, planning, decision making, and procedural and semantic memory.

Information objects in AR environments have unique spatial properties. Because of the nature of gravity, traditional information objects have to be physically attached to the body or other support structures within the environment. However, tools and information objects in AR environments can remain stationary with respect to the world or to user body parts such as the head and the torso and appear to be totally unsupported and floating in space. The amount of mobile space available to organize information objects is increased by extending the working volume from the surface of the body to a peripersonal volume in the volumetric AR computing environment; a working space that is associated with the physical body and, thereby, the user. In such an environment, users will be able to manipulate and access multiple information objects concurrently, much as

users commonly multitask with devices such as cell phones, address books, and other physical information media.

1.4 Research Motivation and Problem Statement

This thesis constructs a spatial framework for information display in AR environments based on experimental behavioral science and neuropsychological studies of how humans interact with visually and physically perceived objects in three-dimensional space. The theoretical framework allows researchers of AR interfaces to design to systematically investigate spatial cognition issues closely related to AR interfaces design.

It seems obvious that the human cognitive system should process information objects in an augmented environment in exactly the same way real information objects are processed. However, information objects in an augmented environment do not necessarily behave the same as objects in reality. For example, tools and information objects in an AR environment can remain stationary in space or be attached to different reference frames in the environment or to body parts. Since it is impossible to generate this apparent “anti-gravity” feature in the physical environment, precious little is known about how humans mentally organize information objects attached to an egocentric or allocentric “weightless” environment. How might users manage and organize different information fields around different frames of reference in this new environment?

The primary attention and efforts for researchers in the AR community has been focused on technologies and engineering of AR systems. User studies in AR are generally limited to testing proof of concept prototypes with simple user evaluation.

Currently there is a lack of explicit theories and guidelines in computer-human interaction to support the design of this emerging technology and its varied applications.

1.5 Contributions of this Thesis

The major contribution of this thesis is the construction of a new spatial framework for information display in AR environments. A large volume of existing work in cognitive psychology and neuroscience is examined and existing theories in human perception and information processing are coalesced and transformed into theories applicable to information placement in an AR environment. Furthermore, six experiments were conducted to discover unique human spatial cognitive properties closely related to the design of AR environments. The experimental results were then combined with existing research in behavioral science and neuropsychology of three-dimensional space and used for the construction of research-based information placement guidelines for mobile AR environments.

The remainder of this dissertation is organized as follow. Chapter 2 reviews literature in behavioral science and neuropsychology that are closely related to spatial information display in AR environments. Chapter 3 presents a spatial framework of three-dimensional spaces based on existing neuropsychological evidence reviewed in Chapter 2. Chapter 4 discusses behavioral properties in the spatial framework based on existing literature. Three research questions are raised and investigated in Chapter 5, 6 and 7. Chapter 5 discusses three experiments that investigate the use of reference frames during the spatial encoding process in AR environments. Chapter 6 discusses 2 experiments to evaluate the applicability of perceptual asymmetry properties in AR interfaces design. Chapter 7 presents a novel metaphor for directing visuo-spatial

attention along with experimental evaluation of the metaphor. The main contributions of this research are then summarized in Chapter 8, and potential future research is discussed.

2 Theoretical Background

Theory driven human-computer interaction design is necessary to develop a high performance AR interface. With motion tracking technologies, AR systems afford many options for information placement relative to the environment, objects in the environment, and the user's body. Figure 2.1 illustrates a prototype AR interface with information attached to different reference frames. If users of AR systems will be accessing, organizing, and deploying large volumes of information in space, then an understanding of how the brain accesses and organizes spatial information is a sound, human factors basis for interface research and guidelines. The problem statement becomes: **given an environment where information can be placed anywhere in space around the user and stabilized relative to the body or the environment, what are effective ways to organize information objects in that space?**

2.1 Spatial Framework of Three-dimensional Space

Much of the cognitive capability of the human brain is allocated to the task of tracking the location of people and objects in space, especially in the planning of motor actions. From biological and psychological viewpoints, AR space is not a continuous Cartesian space. Research in spatial cognition indicates that objects in the environment appear to be modelled in the brain using interrelated spatial coordinate frameworks organized around the body, objects, and the larger environment (Pettigrew and Dreher 1987; Previc 1990b; Bryant 1992; Bryant *et al.* 1992; Pani and Dupree 1994; Cutting and Vishton 1995; Previc 1998).

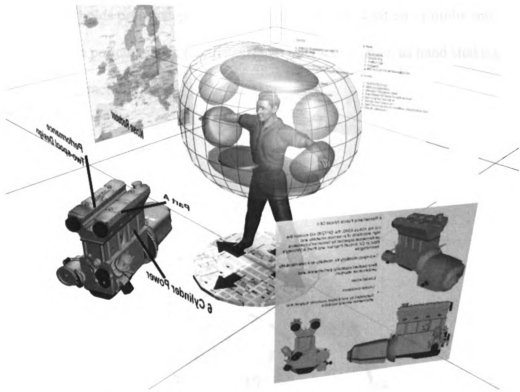


Figure 2.1. A prototype volumetric AR interface with information objects placed in different reference frame.

2.2 Neuropsychology of Three-dimensional Spaces

According to current neuropsychological theories, the brain models the surrounding three-dimensional space as three overlapping regions: (1) personal/body space, (2) peripersonal space, and (3) extrapersonal space.

2.2.1 Personal/Body Space

The clearest psychological spatial boundary is defined by personal space, or body space; it is the psychological space that defines the boundary between the body (the proximal “me”) and the world beyond the body. The personal/body space is the volume extending to a few centimeters from the skin of the body, as illustrated in Figure 2.2.

This space not only holds proprioceptive information about the position of limbs and body; it is also where pericutaneous (tactile surface) interaction (such as hand shaking) and buccal (oral) interactions occur.



Figure 2.2 Personal/body space.

Some neuroscience data based on animal studies suggest that neuronal responses to body space extend slightly beyond the skin surface (Graziano and Gross 1995). Philosophers and psychologists (for example, Heidegger 1968; Bateson 1972) have long speculated that the psychological boundary of the body sometimes expands so that objects near the body are integrated into the personal body space. Although the boundary of the body appears to be physical and fixed from the viewpoint of an objective observer, there is evidence from research in neuropsychology that the sense of the boundary of the body is plastic. Personal space, defined as the shape and extent of body schema, can be expanded to incorporate objects attached to the body (e.g. clothing and tools).

Neuroscience studies by Maravita and Iriki (2004) on neuronal motor responses during tool usage by monkeys suggest that the body schema, defined as receptive fields of neurons associated to perceived body parts, expands to incorporated tools such as sticks and rakes after extended use. Furthermore, they show that this extension of the receptive fields extends to video representations of the monkey's body shown on a monitor, so that the neurons respond to a displaced virtual hand as if it were the monkey's physical hand. This suggests that tools can be incorporated into the body schema at some level.

Another line of research that suggests how media tools can restructure the body schema is work on visual-motor adaptation in space perception. In these studies, a technology is used to alter visual perception through the use of sensory prosthesis such as a prismatic lens. Adaptation to the sensory change, subsequent errors, and readaptation after the alteration is removed are observed (Stratton 1897; Held and Schrank 1959; Harris 1963; Kohler 1964; Hay and Pick 1966; Ebenholtz and Mayer 1968; Dolezal 1982). In studies on visual and motor hand adaptation in virtual environments, it was found that AR systems can remap the perceived location of the hands (motor space) relative to visual space (Rolland, Biocca, Barlow and Kancherla 1995; Biocca and Rolland 1998).

2.2.2 Peripersonal Space

Another key subspace motivated by neuroscience research on three-dimensional spaces is the peripersonal space. Peripersonal space is the volume of space immediately in front of the body and reachable by the arms and hands. Peripersonal space is tied mainly to the egocentric trunk- or shoulder-centered coordinate frame (Previc 1998).

Located immediately in front on the body, biased towards the central 60° in the lower visual field, and with a radial extension of 0-2 m, peripersonal space overlaps considerably with the ergonomic space known as the *reach envelop* (Proctor and Van Zandt 1994a; Proctor and Van Zandt 1994b) (Figure 2.3). Peripersonal space is functionally organized for binocular object inspection, motion processing, hand motion, and manipulations such as directly reaching and handling objects. This interpretation is supported by behavioral evidence, in that information and objects in this area are found and manipulated the fastest (Hari and Jousmäki 1996; Murphy and Goodale 1997).

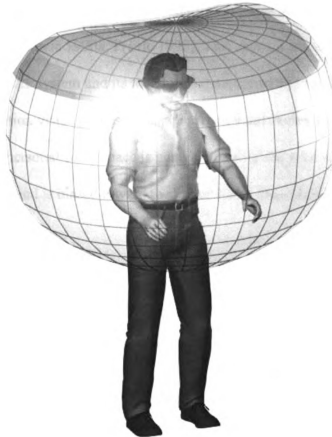


Figure 2.3. Peripersonal space.

2.2.3 *Extrapersonal space*

Extrapersonal space is the spatial volume beyond the reachable distance of the arms. The extrapersonal space can be subdivided into four subspaces: (1) extrapersonal focal space, (2) extrapersonal action space, (3) extrapersonal scene space, and (4) extrapersonal ambient space.

2.2.3.1 *Extrapersonal Focal space*

Extrapersonal focal space is an elliptical region of central fovea vision anchored in the plane of fixation with a lateral extent of 20°-30° and radial extent of higher than 10-20 cm, as illustrated in Figure 2.4 (Rizzolatti, Gentilucci and Matelli 1985; Rizzolatti and Camarda 1987; Previc 1990a; Previc 1998). This space is associated with the retinotopic coordinate system and its location is determined by the fixation of the eyes. It serves high-resolution visual processes that are carried out exclusively in the central visual field. Extrapersonal-focal space is generally associated with visual search and object recognition, and is biased toward the upper visual area slightly outside of reaching distance.



Figure 2.4. Extrapersonal focal space.

2.2.3.2 Extrapersonal Action Space

Extrapersonal action space encapsulates the body in a 360° surround, with a range starting from 2 meters from the body to approximately 30 meters (Figure 2.5). This region appears to be active in orienting and activating attention, memory, and voluntary motor systems within topographically (as opposed to gravitationally) defined external space (Previc 1998), and is biased towards the upper visual field. It is closely linked to the remembrance of specific places or events, in accordance with the general linkage of episodic scene memory to distal space and navigation. It has been argued that the extrapersonal-action space incorporates an allocentric coordinate system, but neuropsychological data and lesion study results provide evidence that the extrapersonal-action space incorporate a gaze-centered or head centered coordinate system.



Figure 2.5. Extrapersonal action space.

2.2.3.3 Scene Space

There is evidence for a mental model of a larger region of visible objects beyond action space. Scene space is not gaze-centered like action space, and involves an allocentrically-oriented model of the larger space around the body (Figure 2.6). This space is assembled from clusters of objects whose position is defined relative to prominent features or objects in a scene (Easton and Sholl 1995; Sholl and Nolin 1997; Shelton and McNamara 2001a; Mou and McNamara 2002). There is evidence of cognitive maps organized and distorted to fit around landmarks and evidence that priming memory for one object activates memory for objects in the cluster or regions nearby (McNamara 1986; McNamara 1989).

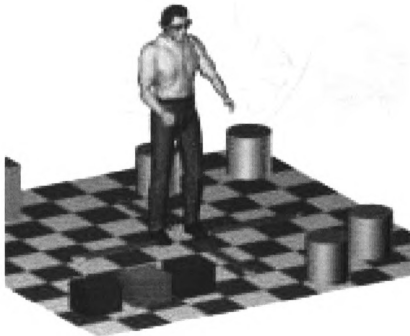


Figure 2.6. Scene space.

2.2.3.4 Extrapersonal Ambient Space

Extrapersonal-ambient space is the outermost space of the visual field (Figure 2.7). It appears to be biased towards the lower visual field. Oriented towards a gravitational, earth-centered spatial framework, it plays a role in the maintenance of spatial orientation, balancing, self-motion (Dictgans and Brandt 1978) and postural control (Previc 1990a; Previc and Neel 1995) and allows the user to interpret self-motion in an apparently stable world (Leibowitz and Post 1982).

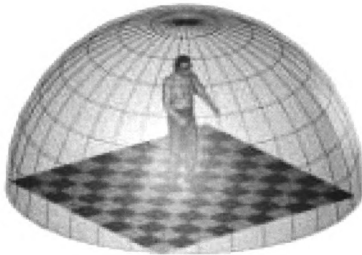


Figure 2.7. Extrapersonal ambient space

2.3 Mapping Digital Information to Space in Augmented Reality Systems

A high performance AR interface design can be constructed by mapping the natural processing properties in different portions of the three-dimensional space to the information placement in the AR environment. In Chapter 3, a spatial information display

framework is constructed based on the literature reviewed in this chapter. Chapter 4 explores the behavioral properties of different portions of three-dimensional space.

3 Spatial Framework of Information Display in Mobile Augmented Reality Environments

A theoretical framework for three dimensional-space based on neuropsychological theories was developed through the examination of existing literature in Chapter 2, segmenting the space around a human in terms of the general use and perception of these spaces. In this chapter, these ideas and other new and existing work will be extended to develop a spatial framework specifically tailored for the presentation of information in mobile AR environments.

3.1 Spatial Information Framework

So, how can the neuropsychological spaces defined in Chapter 2 become information spaces? With motion tracking systems, there are many technological options on how information objects can be placed so as to appear to be stable relative to different reference frames in the spatial framework (Foxlin 2002). Based on the neuropsychological model reviewed in Chapter 2, this chapter establishes a spatial framework for information spaces, which will be referred to as *Infospaces* to emphasize that the spaces are designed to present information.

3.1.1 *Personal-body Infospace*

Information objects attached to the *Personal Body Infospace* remain stationary with respect to some moving part of the body. In order to attach information objects to a moving part of the body, the position and orientation of that body part need to be tracked. Information objects can be attached to any tracked moving body part, such as hands, arms, legs, or other extremities. To examine the possible body-stabilized frames, it is of

interest to examine the skeletal structure of the human body, exploring the major bone groups that can be used to define useful information frames. Figure 3.1 is an illustration of the skeletal structure of the human body. The skeleton represents the rigid structure of the moving elements of the human body and, as such, provides a set of possible tracking references that can define information frames relative to the human body. Since direct attachment to bones for tracking purposes is generally not practical, the attachment is more likely to be to the epidermis (surface of the skin). But proper placement allows the epidural attachment to be a good approximation of the underlying bone tracking.

The concept of a personal body infospace is a very general idea. A human adult skeleton has 206 bones. Clearly, many of these are not useful from an information frame point of view (such as bones in the inner ear) or are redundant (such as the dual function of the ulna and radius or the set of bones in the rib cage). Other bones may have very limited utility in mobile AR environments (such as the bones in the feet). Figure 3.2 describes ten bone groups useful for definition of AR information frames. Some of these groups define frames directly (such as the skull); others define sets of frames (such as the vertebral column).

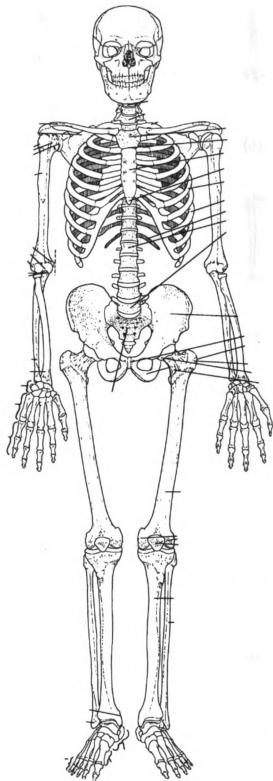


Figure 3.1. Human skeletal structure (Gray, Bannister, Berry and Williams 1995).

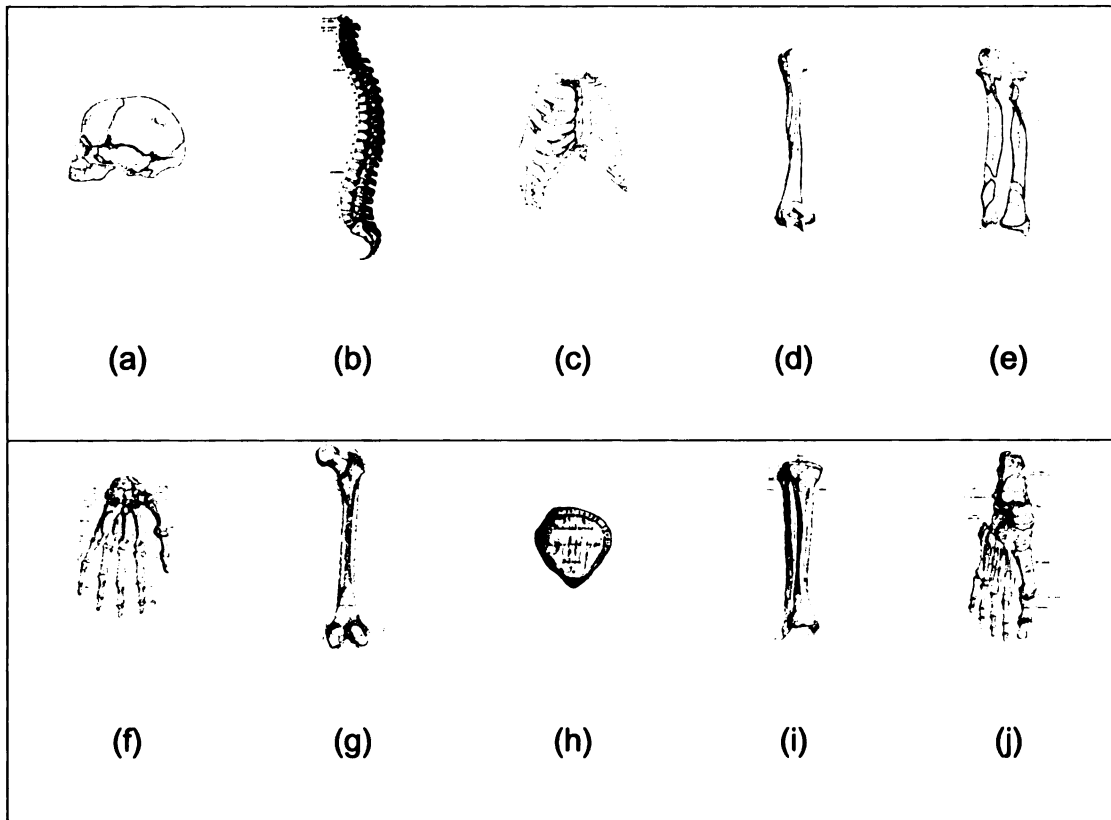


Figure 3.2. Selected bone groups for Personal-Body Infospaces: (a) Skull (b) Vertebral Column (c) Sternum and costal cartilages (d) Humerus (e) Forearm group (f) Hand (g) Femur (h) Patella (i) Leg (j) Foot (Gray et al. 1995).

The Hand Personal-body Infospace is a common infospace for manual interaction in AR environment. Tracking of hand movement is required to facilitate the creation of a Hand Personal-body Infospace. The *human hand* is a complex device with many bones. Figure 3.3 is an illustration of the bones of the human hand. The most basic configuration of tracking would emit 15 frames for a hand, fourteen for the phalanges, the bones of the fingers, and one for the metacarpus. A few technologies exist that can provide this level of tracking for the hand. Simple object manipulation can often be accomplished with only metacarpus tracking.

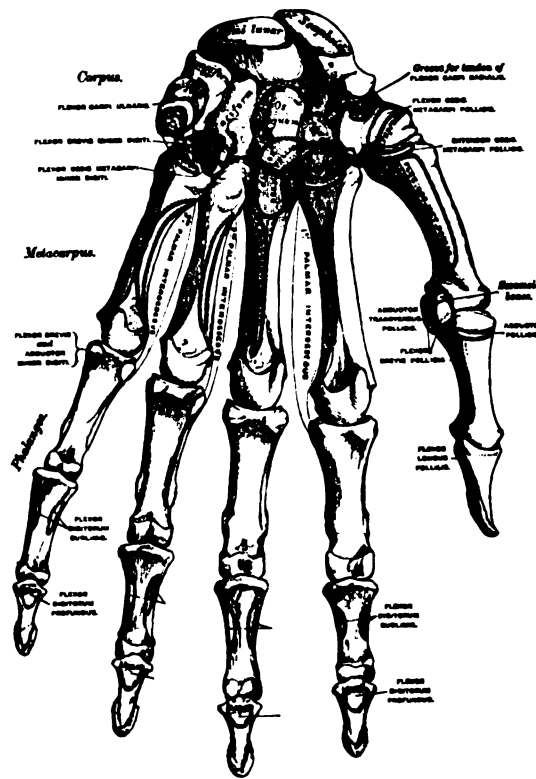


Figure 3.3. Skeletal structure of human hand (Gray et al. 1995).

3.1.2 Peripersonal Infospace

The *Peripersonal Infospace* remains stationary with respect to the upper torso. Tracking of the upper torso is required to create the peripersonal infospace. Tracking of the *sternum* and *costal cartilages* (front part of the body) would generate unwanted breath motion for information objects attached to the peripersonal infospace. Therefore, the *vertebral column* is recommended as the tracking source for the peripersonal infospace, generally through some external attachment such as a belt that will transmit the motion from the vertebral column to a tracking device with a minimum of motion error due to epidermal layers. The Vertebral Column is composed of 7 cervical vertebrae,

12 dorsal vertebrae, 5 lumbar vertebrae, sacrum and coccyx (Figure 3.4). It is situated in the median line of the back of the upper torso. The 7 cervical vertebrae, which form the neck, are not well suited for tracking because of the deformation of muscles around the neck. The 12 dorsal vertebrae and 5 lumbar vertebrae are more suitable for tracking. The vertebral column clearly emits a variety of tracking points, each with unique characteristics. Tracking of the upper back (dorsal area) will create a frame that follows the body.

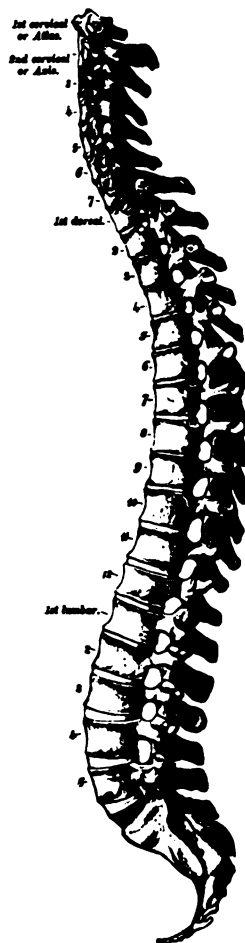


Figure 3.4. The vertebral column (Gray et al. 1995).

3.1.3 *Extrapersonal Focal Infospace*

Information objects attached to the *Extrapersonal Focal Infospace* remain stationary with respect to eye fixation. Tracking of eye movement is required for displaying rendered virtual elements that appear to be stable with respect to eye fixation. When eye tracking is not available, relevant information objects can be placed in a head-stabilized reference frame to grab the user's attention. Position and orientation of the head is commonly tracked in AR systems, typically through tracking of a head-mounted display that is fixed relative to the head.

3.1.4 *Extrapersonal Action-Scene Infospaces*

The *Extrapersonal Action-Scene Infospaces* are an amalgamation of the neuropsychological extrapersonal action space and scene spaces, which define the spatial volume of the allocentrically oriented spaces. Typically, Extrapersonal Action-scene Infospaces encapsulate task-specific working volumes such as desks, cabinets or building structures. Information objects can be attached to stationary objects in the environment without additional tracking support. Some AR systems present information attached to moving objects in the scene. In such AR systems, the position and orientation of objects needs to be tracked. There can be multiple Extrapersonal Action-scene Infospaces existing concurrently for multiple working volumes in a multitasking scenario.

3.1.5 *Extrapersonal Ambient Infospaces*

In current mobile AR systems, information is often presented as “world stabilized”: information is fixed to real world locations in the *world coordinate frame* and the data view varies as the user changes viewpoint orientation and position. Thus, a user's viewpoint position and orientation are tracked and the transformation from the

world frame to the user's view frame is computed and used to transform virtual augmentations and objects so as to appear registered with the real world. In such a system, the only information frame is the world coordinate frame.

3.2 Summary

Based on existing literatures in neuropsychology, a spatial information framework is constructed for information organization for mobile AR computing environment, as summarized in Figure 3.5. The spatial information framework consists of five information spaces, or Infospaces. Chapter 4 reviews a collection of literature about behavioral properties in each Infospaces.

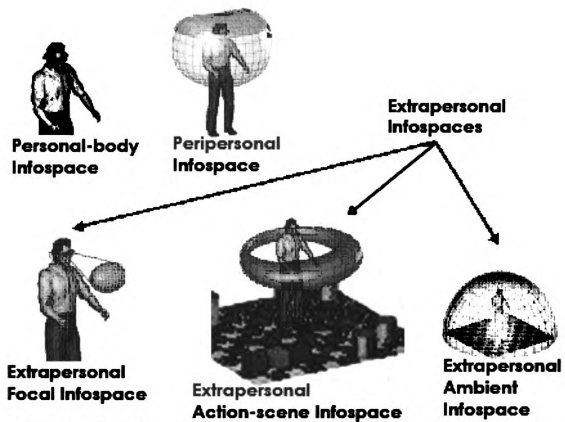


Figure 3.5. Spatial framework for information display in mobile augmented reality environments.

4 Behavioral Properties of Three-dimensional Space

The automatic neurological activities for space around the user can be leveraged for information organization. In order to develop a theory-based interface design, it is important to determine how the human brain perceives and reacts to objects in different spatial location in three-dimensional space. This chapter reviews existing research about behavioral properties of the Infospaces that were defined in Chapter 3, with a mind toward utilizing the spaces for information presentation.

4.1 Behavioral Properties in Personal-body Infospaces

Surfaces on the human body can be used as information spaces. The use of a wristwatch places an information device directly on the surface of the skin. However, beyond this simple functionality and occasional jotting of notes on the skin or decorative uses such as tattoos, physical body space is rarely seen as a possible information space for interaction with information in any form. Tool storage and manipulation using belts or pockets is very common. However, visual displays such as clothing, tattoos, makeup, and other body-attached items are usually not used for communication, particularly with the user himself. They are representations used for signaling information such as social status, sexual availability, and other social information to other observers. Technology now allows the augmentation of the environment with computer-generated virtual content. If a body part can be tracked, a computer can register graphics with the body part and display them in various forms that allow the user to perceive them as placed on or near the body part. This use of the body surface in AR systems to hold and display

private information for the user has considerable potential for providing information to users in a familiar territory, but actual implementations of this idea remain rare.

There are neuropsychological advantages to placing virtual tools such as icons, buttons, and other digital objects on or very near the body surface. Neuroscience data suggests that neuronal responses to body space may extend slightly beyond the skin surface:

... the visual space near the animal is represented as if it were a gelatinous medium surrounding the body that deforms whenever the head rotates or the limbs move. Such a map would divide the location of the visual stimulus with respect to the body surface, in somatotopic coordinates (Graziano et al. 1995/p.1031).

Tracking the position and orientation of individual body parts allows digital information to be attached to the moving body (Owen, Biocca, Tang, Xiao, Mou and Lim 2005). This section reviews some important behavioral properties of the Personal-Body Infospace that are relevant to information communication.

4.1.1 Proprioception

Proprioception is the unconscious perception of movement and orientation of the body arising from sensing mechanisms within the body itself. Neuropsychological literatures suggested that both vision and proprioception contribute to the establishment of spatial representation (Chance, Garnet, Beall and Loomis 1998; Shelton and McNamara 2001b; Yamamoto and Shelton 2005). Literature in the field of

Neuropsychiatry suggests that an essential contribution from *basal ganglia* for the integration of visual and proprioceptive information is required for achieving high accuracy in pointing tasks (Adamovich, Berkinblit, Hening, Sage and Poizner 2001; Keijsers, Admiraal, Cools, Bloem and Gielen 2005). In the design of AR interfaces, additional proprioceptive cues for information objects in the Personal-body Infospace have the potential to increase pointing and manipulation accuracy and naturalness. If information objects are associated with body parts, proprioception assists in the knowledge of the location of the associated object.

In everyday life, human-beings do use the body as a medium for communication, information display, and storage in a limited fashion. The aforementioned watch example is the use of the wrist to display of time and date and for storage of other personal information. Workers attach tools on a waist tool belt for easier and faster access to the tools. The Personal-body Infospace naturally becomes an intuitive information space for metaphorical personalization in AR computing environments. Lehtikainen (2000) proposed the idea of a shirt embedded with an array of pressure sensors that allows the user access to and interactively manipulation of digital information mapped to locations on the torso. Aside from the motor advantage of easy pointing and manipulation, Personal-body Infospaces allow the user to develop metaphorical associations and proprioceptive memory between information objects or control functions and spatial locations on the user's body. Some real life examples of using the body for metaphorical association include the parachute control for skydiving and fishing vests for storing tools on different positions relative to the torso. Metaphorical associations between digital information and control functions to body

positions developed by individual users have a great potential for a more intuitive interface with higher performance. The habituation and proprioceptive memory established could provide a faster and more accurate access, retrieval, and manipulation of information objects and control functions.

4.1.2 Spatial Bias in Personal-body Infospace

Human visual, auditory, and haptic systems for perception and motion are strongly skewed to maximum performance in the ventral (frontal) regions of the body (Corballis and Beale 1983; Corballis 1993). The dorsal regions (back of the body), in general, exhibit decreased sensory resolution, are less accessible by the hands, and are not visible by the user's eyes. Therefore, the back of the body is clearly not ideal for holding digital information that must be viewed or manually manipulated. The Personal-body Infospace is further biased towards the upper body, where the body parts are reachable by the hands.

4.1.3 The Hand and Forearms Personal-body Infospaces

There are cases in current VR and AR practice where information is attached to a limb-stabilized space. In VR interfaces, hand-stabilized information systems often present relevant information such as a crude cursor, virtual representations of the hand, or tools that appear to be attached to or operated by the hand. Information objects attached to a hand-stabilized reference frame should be action orientated. For example, tools selected for the current action, menus, and selection trays can be attached to the non-dominant hand, and the dominant hand can be used for selection and action and manipulation (Figure 4.1). Issues of handedness and kinesthetic asymmetry of the Personal-body Infospace will be discussed in Section 4.4.1.

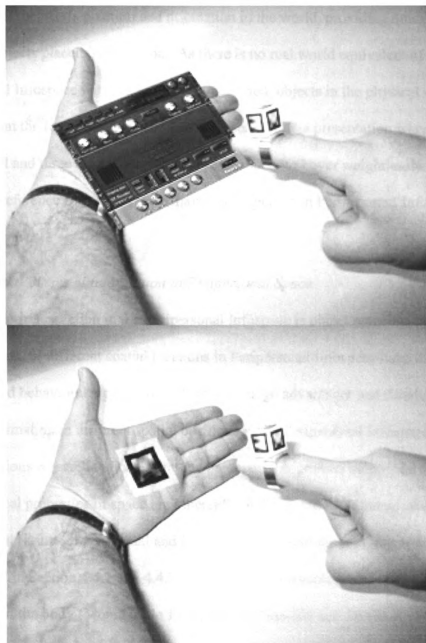


Figure 4.1. Hand Personal-body Infospaces: a menu attached to the non-dominant hand and an interaction tool (the ring) attached to the dominant hand.

4.2 Behavioral Properties in Peripersonal Infospace

Information objects placed in the Peripersonal Infospace remain arm-reachable regardless of the user's position and orientation in the world, providing quick and easy access to objects placed in the frame. As there is no real world equivalent of a Peripersonal Infospace with associated, yet detached, objects in the physical world, experience in the real world does not prepare users for data presentation where two-dimensional and three dimensional data objects appear to hover weightlessly in an egocentric reference frame, and the behavioral properties in Peripersonal Infospace are largely unknown.

4.2.1 Spatial Biases of Information in Peripersonal Space

The primary action in the Peripersonal Infospace is object reaching, grasping and manipulations. If different spatial locations in Peripersonal Infospace have different cognitive and behavioral significance, there are design advantages and disadvantages to placing information in different spatial positions in the Peripersonal Infospace.

Previous research indicates that the perceptual, cognitive, behavioral and biomechanical properties of space are inherently and, sometimes, fundamentally asymmetrical (issues of perceptual and kinematics asymmetries in egocentric spaces will be discussed in Section 4.4.1 and 4.4.2). Reaching movements are biased towards the middle 60° of the body (Mountcastle 1976; Servos, Goodale and Jakobson 1992), the lower visual field and the lower volume of peripersonal space (Previc 1990a; Sheliga, Craighero, Riggio and Rizzolatti 1997).

In an experiment by Biocca, Eastin and Daugherty (2001), it was found that participants were at least 175% and up to 930% faster (average 313%) at locating targets

and placing objects at target locations in the central area of the peripersonal space. Target search and object placement was also found to be significantly faster by 67% within the right side than the left side. A quadrant effect emerged favoring the search, reach, and manipulation for objects in the lower-right quadrant in peripersonal space. This perceptual and motor advantage extended into a memory advantage for recall of the location of objects and recognition for the information objects participants manipulated.

4.3 Behavioral Properties in Extrapersonal Focal Infospace

A key issue in mobile systems is the allocation of spatial attention. The Extrapersonal Focal Infospace is a floating volume centered at the spatial location the user is currently paying attention to. Experimental user interfaces have attempted to harness this high-bandwidth space with eye pointers or “eye mice”. A head stabilized reference frame can be used to display information related to the extrapersonal focal space when eye tracking is not available. However, information objects attached to the head stabilized reference frame remain stationary with respect to the head. This differs from a true eye fixation-based focal infospace, but closely approximates the concept of a space stabilized relative to the vision field.

The main functions of the extrapersonal focal space are object searching and recognition. Objects in the environment are often recognized in the extrapersonal focal space before being brought into the peripersonal or personal-body space. As eye tracking technology inside AR environments is technologically impractical at this time, this section will focus on important behavioral properties of a head stabilized reference frame and human spatial attention.

4.3.1 Visual Clutter in Head Stabilized Reference Frame

Head-stabilized reference frames are the most common space for presenting non-task-related information in AR and wearable computing environments. No tracking is necessary for the presentation of head-stabilized data. There is extensive research in displaying information for drivers and pilots thru Head-up Displays (HUDs) and HMDs. A study conducted by Haines, *et al.* (Haines, Fischer and Price 1980) indicated that pilots who use an HUD have less head and eye movement when compared to pilots that use traditional displays in the cockpit panels. However, several reports indicate that optically overlaid information cannot be processed in parallel (Neisser and Becklen 1975; Becklen and Cervone 1983; McCann, Foyle and Johnston 1994). Others have reported that there is a reaction latency associated with cognitive switching among the environment and the overlaid information (Fisher, Haines and Price 1980; Weintraub, Haines and Randle 1985; Larish and Wickens 1991), and symbology placed within a 5 degree radius of the fovea is annoying to drivers (Sojourner and Antin 1990; Inzuka, Osumi and Shinkai 1991). These research results suggest that only a small amount of information can be placed in the Extrapersonal Focal Infospace, and the central visual field should be reserved to avoid visual clutter to the real environment.

4.3.2 Perceptual Fading of Visual Stimulus

It is well known that perception of sustained and constant sensory input attenuate over a period of time ranging from seconds to minutes. This “perceptual fading effect” has been well documented in perceptual psychology in vision (Ditchburn and Ginsborg 1952; Riggs and Ratliff 1952; Riggs, Ratliff, Cornsweet and Cornsweet 1953; Krauskopf and Riggs 1959; Heckenmueller 1965), audition (Hood 1950), touch (Hoagland 1933),

smell (Engen 1982) and taste (Abrahams, Krakauer and Dallenbach 1937). One everyday example of this phenomenon for the visual sensory channel is that dirt particles on eyeglasses will perceptually disappear after a few seconds if the person is not intentionally paying attention to it. While information attached to the head stabilized reference frame is always visible to the user, interface designers need to be aware that these information objects could perceptually disappear over a period of time ranging from seconds to minutes.

4.3.3 Spatial Bias in Head Stabilized Reference Frame

Human visual attention is biased towards the central area of the head stabilized reference frame. However, the central area should be reserved to avoid visual clutter to the real environment and to avoid degradation of navigation. It is suggested that information objects be placed at the peripheral area of a head stabilized reference frame (McCann *et al.* 1994).

4.4 Behavioral Properties in Relation to Egocentric Infospaces

An Egocentric Infospace is registered with and moves with some part of the body. This attachment personalizes the space for a given user and allows the space to follow the user or specific user appendages. User interface elements residing in an egocentric Infospace appear to be attached to the user, either directly or through some invisible attachment. The Personal-body Infospace, Peripersonal Infospace and Extrapersonal Focal Infospace are all egocentric Infospaces. An interesting characteristic of egocentric Infospaces is the existence of spatial biases due to asymmetries of the brain and body that effect perception of and interaction with user interface components in this very personal space around the body.

Psychological research has demonstrated that human behavior consistently exhibits egocentric spatial biases. There are well understood perceptual asymmetries in psychology and neuroscience for the left/right, upper/lower, and near/far visual fields. Motor actions are also highly asymmetric due to handedness. These asymmetries influence the perception and interaction of information at various spatial locations.

Placing an object at different locations could significantly alter the cognitive process. These effects are relatively benign for traditional user interfaces due to fixed placement and layout of physical interface components (i.e. display, keyboards and the mouse). Due to the limited field of view of small display devices and the fact that information objects are usually attached to allocentric reference frames, the spatial locations either do not consistently stay on one side of any of the known zones of asymmetry or varying placement is not an option at all due to limited screen size.

Egocentric Infospaces allow placement of information that moves in a manner directly related to body elements. Hence, placement of interface elements can be managed in relation to known spatial asymmetries, allowing spatially significant regions around the body to be exploited in egocentric Infospaces.

4.4.1 Kinematics Asymmetry

One of the advantages of immersive AR interfaces is that users can apply intuitive bimanual interaction, the use of both hands in interface tasks. Bimanual interaction in conventional user interfaces is generally limited to the hand-cooperative task of typing. It is well-known that human motor skills are asymmetric due to handedness and cerebral lateralization. This thesis will examine issues of kinematic asymmetry in the context of bimanual action.

Unimanual tasks, tasks that can be completed by one hand, are usually biased towards the dominant hand. So it is a simple design guideline that the system should present simple pointing and selection tasks on the side of the user's dominant hand. The dominant hand is excellent at precise, corrective and rapid movements, while the non-dominant hand usually acts in a supporting role or as a frame of reference for the dominant hand.

Guiard's Kinematic Chain Theory (Guiard 1987; Guiard and Ferrand 1995) provides a theoretical framework for the role of the hands in bimanual activities, and how the actions of the two hands work complement each other. The theory classifies bimanual asymmetric actions in the following three classes:

1. Spatial Reference in Manual Motion: motion of the dominant hand is often based on a spatial reference defined by motion of the non-dominant hand. The roles of the non-dominant hand include a physical stabilizing action (e.g. stabilizing the paper when writing), defining steady states (e.g. putting the non-dominant hand in front when hitting a tennis ball with a racket using the dominant hand), or defining a spatial reference. Bimanual tasks that involve information objects requiring a physical stabilizing action, defined steady states, or a defining spatial reference should be placed on the non-dominant side.

2. Contrast in the Spatial-Temporal Scale of Motion: the dominant hand has a considerably finer spatial and temporal motor resolution. Information objects for bimanual interactions that require macrometric movement should be presented to the side of the non-dominant hand, and tasks that require micrometric movement should be presented to the side of the dominant hand.

3. *Precedence in Action*: the non-dominant hand is typically the first participant in bimanual interaction, with motion preceding that of the dominant hand. A single interface element that requires bimanual interactions should be presented on the side of the non-dominant hand

Guiard's Kinematic Chain Theory provides a framework for the design of user interface components based on bimanual computer interaction. Interchanging subtasks between the two hands can potentially impact the performance of a task and circumstances in the design that encourage this interaction should be avoided. Presenting information on the correct side of the body would result in a faster and more natural access to the relevant information objects by the hands.

Tasks that are naturally designed for bimanual interaction are best placed on the non-dominant side of the body so as to encourage reach and acquisition by the non-dominant hand rather than an acquisition by the dominant hand that may force a transfer to the non-dominant hand for stabilization or referencing. Since the non-dominant hand is typically the initiator in a bimanual interaction, forcing reach on the dominant side by placement can delay the onset of bimanual interaction and, again, force a transfer.

4.4.2 *Perceptual Asymmetries*

Perceptual asymmetries refer to the asymmetric properties of human perception in different visual fields. It is well known in psychology that humans perceive the same stimulus material differently when it is presented in different spatial locations. The location of an object affects cognitive processes in the brain.

4.4.2.1 Bilateral Asymmetry

The concept of contralaterality (the difference in information processing between two sides of the brain) was documented as early as 2500 B.C.E. by the ancient Egyptians (Hecaen and Albert 1978). The human mind consistently exhibits a left-right bias in visual perception and information processing. Information in the left visual field *of both eyes* is sensed by the right side of the retinas and then transmitter over the visual pathway leading to the visual cortex of the right hemisphere. Similarly, information in the right visual field is sensed by the left side of the retinas and then transmitted over the visual pathway leading to the visual cortex of the left hemisphere. Figure 4.2 illustrates these visual pathways. Hence, visual information from the left and right visual fields projects exclusively to the contralateral cerebral hemispheres (Bryden 1982).

Research in visual perception is often based on the use of conventional tachistoscopic techniques to test hypotheses relative to perceptual bilateral asymmetry effects. A subject is asked to fixate on the center of a screen, and stimulus materials are flashed to either the left or right side of the visual field. Reaction time and/or task accuracy are measured. The reaction time differential measured in this experimental technique is very short. Typically, 100 milliseconds is considered a significant effect (Solso 1998). To summarize various experimental results undertaken by various researchers using the tachistoscopic technique and lesion studies, the left hemisphere (right visual field) is found to be biased towards letters and words, language (Strauss 1998), functional or symbolic meaning, verbal memory, local patterns (Robertson and Lamb 1991; Yovel, Yovel and Levy 2001), higher spatial frequencies (Sergent 1983; Sergent 1987), categorical spatial relationships (Kosslyn 1987), and time; while the right

hemisphere (left visual field) is found to be biased towards geometric patterns, visual appearance, visual memory, global patterns, lower spatial frequency, coordinate spatial relations, emotion (Dimond and Farrington 1977), face recognition, and sustained attention (Whitehead 1991). These experimental results are summarized in Table 4.1.

Left hemisphere (Right visual field)	Right hemisphere (Left visual field)
Letter and words	Geometric patterns
Functional or symbolic meaning	Visual appearance
Verbal memory	Visual memory
Local patterns	Global patterns
High spatial frequencies	Low spatial frequencies
Categorical spatial relations	Coordinate spatial relations
Time	Emotion
	Face recognition
	Sustained attention

Table 4.1. Summary of cerebral hemispheric specializations.

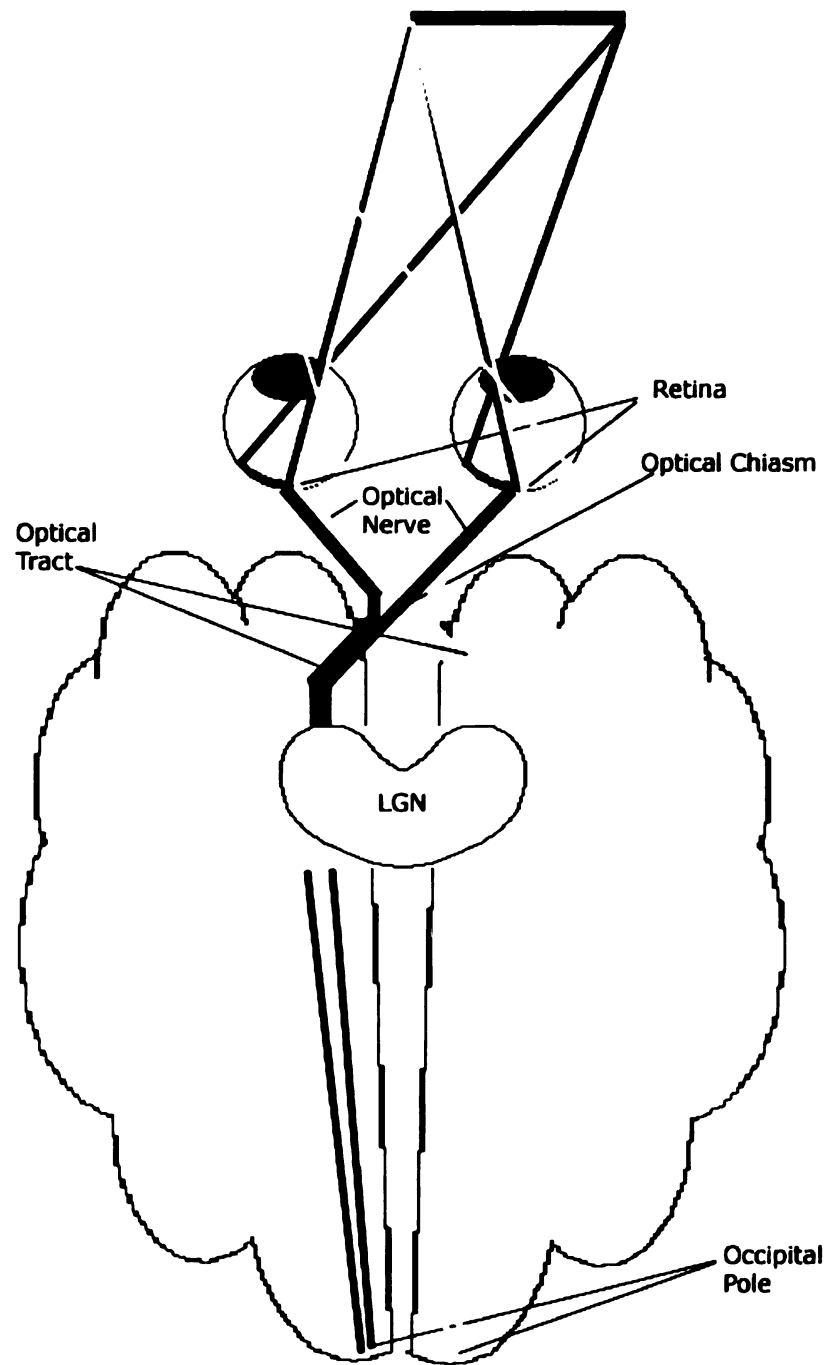


Figure 4.2. The visual pathway of left and right visual field. Retinal signal from the left and right visual fields projects exclusively to the contralateral cerebral hemispheres.

4.4.2.2 Perceptual Asymmetry: Upper vs. Lower Visual Field

Besides the well-known left and right hemispheric specialization, there are also perceptual and behavioral asymmetries within the upper/lower visual field and the far/near visual field. The optic nerves direct the retinal signals to the optical chiasm, where retinal signals from the left and right visual field are divided and fed to the contralateral hemisphere as shown in Figure 4.2. Figure 4.3 illustrates the further processing of visual information in the brain. After the optical chiasm, retinal signals proceed through the optical tract, to the lateral geniculate nucleus (LGN). From there, the visual pathway of the upper and lower visual quadrant spilt into two routes, where the pathway of the upper visual quadrant takes a longer route to the temporal lobe (this pathway is known as the Meyer's Loop) before it heads to the occipital lobe. The representation of the upper and lower visual quadrants is anatomically discontinued in the extrastriate areas (V2 and higher), with visual area V1 physically separating them (Rubin, Nakayama and Shapley 1996).

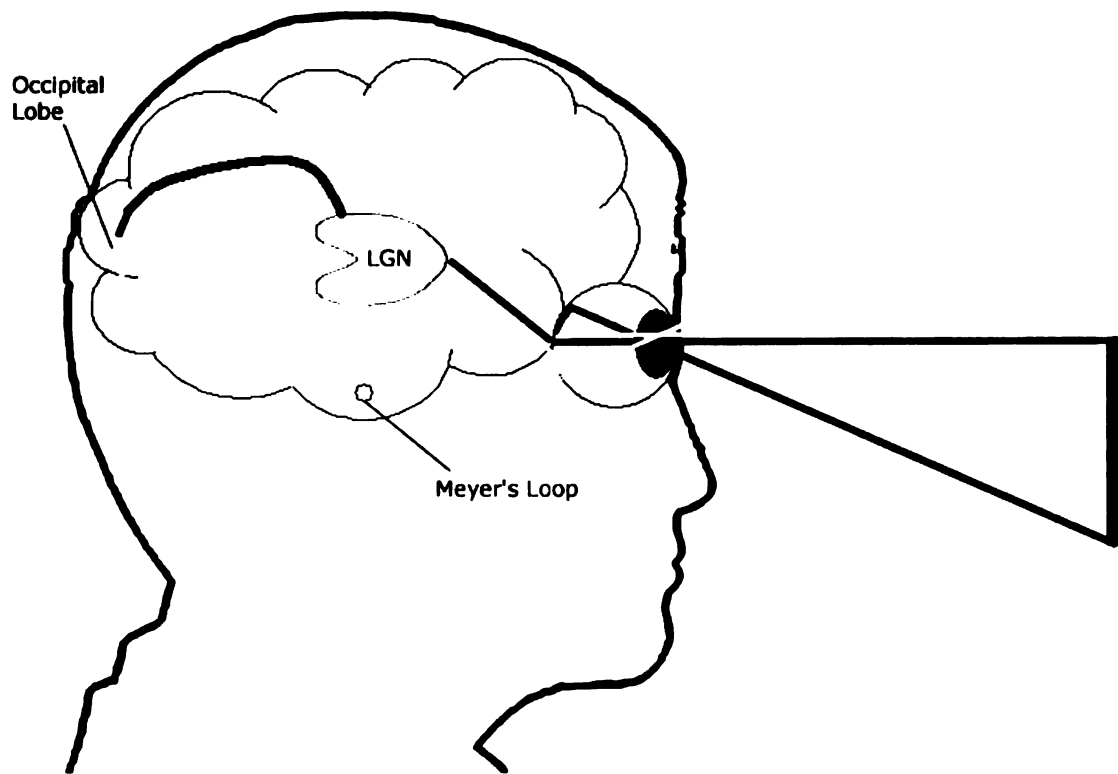


Figure 4.3. Visual pathway of the upper and lower visual field.

Previc (Previc 1990a) argued that most reaching and grasping behavior occurs in the lower visual field whereas visual search and object recognition occur in the upper visual field. So the lower visual field became specialized for reaching and other visuomotor activities and the upper visual field became specialized for visual search and object recognition (Previc and Blume 1993). Previc also argued that visual attention can be subdivided into two major systems: (1) a peripersonal system that assists in reaching and other visuomotor activities, which is biased towards the lower visual field; and (2) an extrapersonal system that is used in visual search and scanning, which is biased towards the upper visual field. Objects in the near space are primarily processed by visuomotor

systems for reaching and grasping, whereas objects far from an observer are primarily processed with visual search and object recognition.

An upper-field bias in visual search/scanning has been shown in studies that used single-fixation search field presentation and those that allowed free eye-movement search. According to these results, the upper visual field is better for presenting dynamic information, which requires object recognition and visual search, whereas the lower visual field is better for presenting static information that has already been recognized or that requires visuomotor activities. Experimental results from studies conducted by Wade, *et al.* demonstrated that the lower visual field is specialized for perceiving shape from shading information whereas the upper visual field is specialized for perceiving shape from edge-based information (Wada, Saijo and Kato 1998). Rubin, *et al.* (Rubin *et al.* 1996) also found that the lower visual field performed much better than the upper visual field in the segmentation of an image into figures and backgrounds. Previc (Previc *et al.* 1993) investigated visual search performance as a function of a target's location in space. The ability to find a target shape was best when it was presented in the upper-right visual field and was closest to the fixation point in its depth and eccentricity. This result is consistent with the individual results for spatial subdivision.

4.5 Behavioral Properties in Extrapersonal Action-scene Infospaces

In most current AR systems, information is presented in Extrapersonal Action-scene Infospaces, i.e., objects appear to be stationary relative to the surrounding environment. In this scenario, virtual objects provide task-specific augmentation or online information about the real environment (Caudell and Mizell 1992; Feiner, MacIntyre and Seligmann 1993; Feiner, MacIntyre, Tobias and Webster 1997; Tang,

Owen, Biocca and Mou 2003). These systems require proper registration between the real and the virtual environment so the virtual object appears to be stationary and in the correct location corresponding to the real environment. By spatially relating virtual information to physical objects and locations in the real world, AR provides the human cognitive system with strong additional leverage in many tasks.

4.5.1 Spatial Consistency of Information Objects with the Environment

By “seaming” the information to the real environment, AR technologies are used “as a complement of human cognitive processes” (Neumann and Majoros 1998). There is evidence that the cognitive load for processing virtual information objects can be reduced when information objects are spatially consistent with the environment. The cost for information search and attention switching between a workpiece and detached media (such as a paper manual) can be reduced by spatially placing task related information in the correct spatial location. Tang *et al.* (Tang *et al.* 2003) designed and evaluated an AR system with spatially registered three-dimensional instructions that directs the user during an assembly process using spatially registered instructions stabilized to the workpiece. Experimental data demonstrated that subjects using the AR system achieved a lower error rate and perceived a lower mental effort compared with subjects using other traditional instructional media (paper and conventional screen-based presentations, for example). The experimental results seem to indicate that the cognitive system processes spatial information and operations (e.g. mental rotation, spatial memory and spatial updating) of virtual objects along with real objects and the environment, and, as a result, spatially registered instruction presentation relieves the mental effort for processing spatial information for the virtual objects. Psotka (Psotka) conducted an experiment to evaluate

visual memory of pictures in three conditions: (1) a *Monitor Condition* with the pictures being displayed in a stationary monitor, (2) a *Virtual Reality Condition* with the pictures floating in the air around the user in a virtual environment, and (3) an *Augmented Virtual Reality Condition* with the pictures projected on the physical wall of the experiment room. The results show that subjects in the Augmented Virtual Reality Condition recalled twice as many items as those recalled by subjects utilizing either of the other two conditions. The author interpreted the increased memory effect as a result of spatial consistency of virtual objects to the real-world coordinate system.

4.5.2 Remote Interaction for Information Objects in Extrapersonal Action-scene

Infospace: Selection and Manipulation

Information objects in an Extrapersonal Action-scene Infospace are attached to allocentric reference frames. Their visibility and reachability are dependent upon the user's location and orientation. Thus, information objects may fall outside the reachable distance of user's hands as the user navigates in the environment. In some cases, information objects can be reachable by navigating towards the objects so that they are within reachable range, while in other cases, it would be more convenient to interact with the information remotely.

There are two types of remote interaction: selection and manipulation. Selection is choosing an object in three-dimensional space. Manipulation involves the selection of an object, modifying its position and orientation, and then releasing it. There have been many studies in VR and traditional human-computer interaction that explore different options in object selection and manipulation. This section explores the studies relevant to, or leading to, ideas on selection and manipulation in AR user interfaces.

4.5.2.1 Remote Object Selection

Different body parts (such as the finger, hand, and head) can be used as interface elements for selection of remote objects. A measurement of difficulty of input devices for a pointing task can be calculated using Fitts's Law (Fitts 1954; Fitts and Peterson 1964). Fitts's Law states that the time required to complete a pointing task is *directly proportional* to the distance to the target, and *inversely proportional* to the width of the target. In other words, the closer and/or the larger the target, the shorter the time required for the pointing task. Fitts's Law has been shown to be valid under a wide variety of circumstances, including movement of different body parts (such as fingers, hands, arms, foot, head, and eye-gaze), pointing tasks of different pointing devices (such as mouse, joystick, touch pad, trackball, touch screen), varying physical environments (such as underwater), and diverse user populations (such children, aged, different gender).

Fitts's Law can be used to compare and contrast performance of the same pointing task using different input mechanisms. Experimental data by Langolf (Langolf 1973) shows that performance of a pointing task using the head has the highest Fitts's Index of Difficulty (i.e. most difficult) followed by the arm, the hand, and then the finger. In other words, pointing using the head is very difficult, both in terms of precision and time of completion, followed by the arm, the hand, and then the finger. This result suggests that body parts used for pointing and selection tasks should be chosen based on the order of finger, hand, arm, and then head.

4.5.2.2 Remote Object Manipulation

A few studies in VR interaction techniques have been conducted to explore techniques to manipulate virtual objects outside the reachable distance of the arms. In the

Ray-casting technique, a user utilizes a virtual light that extends from the hand to select and manipulate objects. With the Ray-casting technique, the user aims at the target object using this virtual beam of light. When an appropriate target object has been selected, usually through pressing a button held in the hand, the target object is attached to the virtual light, and spatial location and orientation can be manipulated with simple hand movements. The object is then released, usually due to a button release or secondary button press event and the manipulation is complete. Ray-casting is a highly intuitive technique. However, manipulation using the Ray-casting technique exhibits a “lever-arm problem”: the selected object is attached to the end of a long lever arm, making distance manipulation and arbitrary rotational control of the object impossible.

Mine (Mine 1996) developed the *CHIMP (Chapel Hill Immersive Modeling Program) interaction technique* for distant objects in virtual environments. Similar to the Ray-casting technique, Mine’s remote manipulation technique uses a spotlight attached to the hand for target object selection. Once the object is selected for manipulation, a pop-up menu appears so the user can specify the translation, rotation and other manipulation values through a pop-up keyboard or scroll-bars. While this technique allows the user to manipulate target object accurately by entering precise numeric manipulation values, it is very unintuitive and unnatural.

The *Arm-extension technique*, or the *Go-Go technique* (Poupyrev, Billinghurst, Weghorst and Ichikawa 1996), incorporates an extendable virtual hand such that a user can grab remote objects in the virtual environment. The Go-Go metaphor is implemented by using a nonlinear function for mapping the movement of the user’s physical hand to the movement of the virtual hand. The user reaches the target object by extending the

physical hand toward the object of interest, and the virtual hand extends to a remote location through the mapping of the non-linear function. This technique allows full six degree of freedom manipulation of target objects, and manipulation is very intuitive once the target object is selected. However, object selection is difficult because of the non-linear mapping of the virtual hand, overriding proprioceptive cues and making the control of movement of the virtual hand very difficult.

The *World in Miniature Technique* (Stoakley, Conway and Pausch 1995) provides a miniature replica of the world where the user can manipulate all the objects in the replica within reachable distance. Manipulation of objects in the miniature world maps to full scale manipulation of objects in the actual world. This technique allows users to manipulate all objects freely regardless of the user's location and orientation. However, it is hard to accurately perform micro-manipulations in the miniature world when the scaling from the miniature world to the full-sized world is large. Furthermore, there is a mental effort overhead for performing spatial transformation between the miniature world and the full-scale world.

In the *HOMER (Hand-centered Object Manipulation Extending Ray-casting)* technique (Bowman and Hodges 1997), a user first selects the target object as in the Ray-casting technique. Once the target object is selected, orientation of the object is controlled by the orientation of the user's hand, and the position of the object is controlled by the position between the user's hand and user's body. Since there are limits on the range of hand position and orientation changes within the reach envelope, this method, while intuitive, can only effect limited changes in object orientation and position.

In the *Voodoo Doll Technique* (Pierce, Stearns and Pausch 1999), a user first selects the target object as in the Ray-casting technique. Once the target object is selected, a replica of that object appears in front of the user and manipulation of the replicated object results in a corresponding manipulation of the remote target object. This technique achieves a rather satisfactory accuracy and control in orientation manipulation, but the direction and scale of position manipulation is not clear to the user during the manipulation. An improved Voodoo Doll Technique (Pierce and Pausch 2002) was created by adding a reference point to both the target object and replicated object after selection. There are reports indicating that users are occasionally confused as to which of two objects is the voodoo doll.

There are advantages and disadvantages for every manipulation technique. Characteristics of each manipulation technique that illustrate these advantages and disadvantages are summarized in Table 4.2. There is no standard interaction technique that will work for all applications. AR interface designers will need to analyze the requirements for a specific application before choosing a manipulation technique.

Manipulation Method		
Ray-casting	Pros	Extremely Intuitive
	Cons	Limited freedom in position and orientation manipulation
CHIMP	Pros	Precise manipulation
	Cons	Magnitude input is cumbersome
Arm-extension	Pros	Intuitive Manipulation
	Cons	Object selection and position manipulation is hard to control
World in miniature	Pros	Manipulation of all objects at any time
	Cons	Low accuracy in position manipulation Mental effort overhead for spatial transformation
HOMER	Pros	Relatively intuitive
	Cons	Limited freedom in orientation manipulation
Voodoo Doll	Pros	Relatively intuitive
	Cons	Occasional confusion between the control doll and the target object

Table 4.2. Summary of pros and cons of different remote objects manipulation method

4.5.3 *Unregistered Extrapersonal Action-scene Infospace*

Extrapersonal Action-scene Infospaces do not necessarily need to be registered with the real environment. Unregistered Extrapersonal Action-scene Infospaces are spatially independent of the real environment. The volume of Extrapersonal Action-scene Infospaces is much larger than the egocentric reference frames, and is extendable without a limit. Unregistered Extrapersonal Action-scene Infospaces are suitable for applications with large volumes of information objects. It can be used as a working volume for browsing, searching, and management of non-task-specific data.

4.6 *Behavioral Properties in Extrapersonal Ambient Infospace*

Extrapersonal-ambient space is the outermost space of the visual field. In the human cognitive system, this space is primarily used for motion perception, maintaining spatial orientation and postural control. Extrapersonal Ambient Infospace is not commonly used for displaying digital information or user interface design. In a real world scenario, extrapersonal ambient space is often the conveyer of implicit information such as time of the day as indicated by the status of the sun or moon in the sky and relative location in space as indicated by landmarks. Information in this space is Earth-fixed and generally not task or object specific.

4.6.1 *Spatial Bias in Extrapersonal Ambient Infospace*

Extrapersonal Ambient Infospace is biased towards the peripheral visual field (Dictgans *et al.* 1978; Leibowitz *et al.* 1982; Previc *et al.* 1995). Information in a person's peripheral vision is usually processed without conscious attention. The Extrapersonal Ambient Infospace is also biased towards the lower visual field (Foley and

McChesney 1976; Telford and Frost 1993; D'Avossa and Kersten 1996) for the perception of vection and optical flow on the ground during forward locomotion.

4.6.2 Linear Perspective and Motion Perception Properties

The visual cues that are the most significant in the extrapersonal ambient space are those related to motion perception and spatial orientation, such as horizontality cues, linear perspective, and optical flow. Extrapersonal ambient space is particularly sensitive to motion information in all three types of angular motion (yaw, pitch and roll), as well as inward linear motion (Wallach 1987). Extrapersonal ambient space is less sensitive to linear motion moving outward, side to side, and up and down. The evolutionary or developmental explanation for this property is that human beings rarely walk backward, side to side, or up and down. So the human brain was evolved or developed to be specialized in one type of linear motion visual processing.

4.7 Summary

The literature reviewed provided a solid basis for mapping spatial cognitive properties of different Infospaces for the design of AR environments. At the same time the reviewed literature show a gap in research about other unexplored cognitive properties in spatial framework that is useful for the design of AR environments. Chapter 5, 6 and 7 present three sets of experiments that address important research questions in relation to spatial frameworks. The answers to these questions can be use to optimize spatial placement of information in AR environments.

5 Reference Frames in Mobile Augmented Reality Displays

The first question about behavioral properties in Peripersonal Infospace is: Can the human cognition system manage to process information objects attached to the egocentric reference frame naturally? For example, when users turn left with their eyes closed, will the mind's information mapping assume a surrounding information array in an egocentric frame will move with the body or will the cognitive systems assume the objects will stay still with respect to the world? Is this cognitive behavior fixed or does it adapt in the presence of new information display techniques? *Can the human cognition system process information objects attached to an egocentric reference frame?*

It is clear that each frame of reference has its own advantage in some applications. However, it is not clear how to manipulate a user's preference of frames of reference according to different applications. Three experiments were conducted to investigate the default reference frame for spatial memory, and how to manipulate human spatial cognitive systems to adapt to a different reference frame (Mou, Biocca, Owen, Tang, Xiao and Lim 2004a).

5.1 Related Works

This thesis presents the first specific research into human spatial memory and spatial updating of “weightless” information array in AR systems. However, some answers to the above questions may be suggested by human spatial memory and spatial updating of real objects in the physical world. There is a large body of evidence indicating that human spatial cognition updates locations of objects during locomotion (for example, Levine, Jankovic and Palij 1982; Rieser, Guth and Hill 1986; Rieser 1989;

Presson and Montello 1994; Farrell and Robertson 1998; Simons and Wang 1998; Wang and Simons 1999; Sholl and Bartels 2002; Waller, Montello, Richardson and Hegarty 2002; Mou, McNamara, Valiquette and Rump 2004b; Mou, Zhang and McNamara 2004c). For example, participants in one of Waller et al.'s (2002) experiments learned 4-point paths. In the "stay" condition, participants remained at the study position and made pointing judgments from headings of 0° and 180° ("aligned" vs. "misaligned"). The results in this condition replicated several other studies of spatial memory in showing that performance was better for the imagined heading of 0° than for the imagined heading of 180° (e.g. Levine *et al.* 1982). In the "rotate–update" condition, participants learned the layout and then were told to turn 180° in place so that the path was behind them. Performance was now better for the heading of 180° (the new egocentric heading) than for the heading of 0° (the original learning heading). This result indicated that, as they turned, participants updated their orientation with respect to the locations in memory.

Simons and (see Simons *et al.* 1998; Wang *et al.* 1999) investigated the interaction between observer movement and layout rotations on change detection. They showed that detection of changes to a recently viewed layout of objects was disrupted when the layout was rotated to a new view and the observer remained stationary, but there was no disruption when the layout remained stationary and the observer moved to the new viewpoint. In other words, updating was efficient when the observer moved around the layout but not when the layout rotated in front of the observer (see Wraga, Creem and Proffitt 2000, for analogous results in imagined updating).

In a recent study, Mou, McNamara, et al. (2004b) reported that the angular distance between both the imagined heading and the learning heading and the imagined

heading and the actual heading had effects on people's ability to accurately point to objects in the environment. Participants in one of their experiments learned the locations of 10 objects from a single view (e.g., a vase was located next to the learning position; see Figure 7 of that article), walked to the center of the layout (e.g., next to a shoe), and faced three headings before making pointing judgments from imagined headings. There were three imagined headings, 0° (e.g., "Imagine you are facing the phone), 90° ("Imagine you are facing the banana"), or 225° ("Imagine you are facing the jar"), and there were two angular distances between the imagined heading and the actual heading, 0° (e.g., participants actually faced the phone and were instructed to imagine facing the phone) or 225° (e.g., participants actually faced the book and were instructed to imagine facing the phone). Pointing performance was best when the imagined heading was parallel to the learning view. Pointing performance was also better when the actual and the imagined headings were the same. Mou, McNamara, et al. proposed that people both represent locations of objects in terms of an object-to-object frame of reference selected by the egocentric view (also see Mou *et al.* 2002; Mou *et al.* 2004c) and update their location and orientation in terms of that frame of reference during locomotion.

In Experiment 1, using the paradigm developed by Mou, McNamara, et al. (2004b), we investigated whether people with no experience in mobile AR systems would use the environment-stabilized or body-stabilized frame of reference as the default. We hypothesized that if participants used the body-stabilized frame of reference, the angular distance between the imagined heading and the actual heading would not affect pointing performance; however, if they used the environment-stabilized frame of reference, the angular distance between the imagined heading and the actual heading

would affect pointing performance, just as was observed in the Mou, McNamara, et al. study. We only investigated the user's frame of reference preference during rotation (and only in the horizontal plane) rather than in translation (in all three body axes); we assumed the information objects around the user's body should be arrayed independently of the user's translation. We limited our study to the frame of reference preference during body rotation rather than head rotation because a display stabilized with respect to the head would have a very limited information field. In this study, the second goal was to examine whether the nature of the representation of the objects in the AR system can be altered from environment centered to body centered. The experience of large objects moving with the body does not occur normally in the real world, except in cases in which objects are directly attached to the body. So, although the default organization of virtual objects appears to be tied to the exocentric world frame, experience of a body-stabilized frame might enable users to adopt the newly experienced frame when updating their memories for objects' locations in a new layout, even without direct visual guidance. In Experiment 2, we examined whether a couple of minutes of experience in the body-stabilized AR display would allow users to adopt the body-stabilized frame of reference. In Experiment 3, we examine whether only oral instructions to use a body-stabilized frame of reference for updating the location of a set of objects might be sufficient to induce participants to use a body-stabilized frame of reference in accessing a complex layout. Waller et al. (2002) reported that people were able to imagine simple, body-stabilized 4-point paths in front of them when they physically turned back after being instructed to do so.

5.2 Experiment 1: The Default Reference Frame

In Experiment 1, participants learned the locations of virtual objects displayed on the floor from a single stationary viewing position in a large cylindrical room; they were instructed either to

5.2.1 Methodology

5.2.1.1 Participants

Participants were 8 female and 8 male undergraduates at Michigan State University who participated voluntarily as partial fulfillment of course requirements.

5.2.1.2 Materials and Design

Stimulus materials were displayed in stereo with the Sony Glasstron LDI-100B head mounted display. Head motion was tracked with a Polhemus Fastrak magnetic tracker. Stereo graphics were rendered in real time on the basis of the data from the tracker. Presentation of stimulus materials, audio instructions for participants, experimental procedure sequencing, and data collection for the experiment were automated so that the experimenter did not need to hand code the experimental results. The experiment was developed using the ImageTclAR augmented reality development environment (Owen, Tang and Xiao 2003).

The list of objects used in the experiment is illustrated in Figure 5.1. The configuration of eight virtual objects was displayed by the AR system (see Figure 5.2). Objects were selected with the restrictions that they be visually distinct, fit within an area approximately 0.3 m on each side, and not share any obvious semantic associations. The

objects were all virtual analogs of existing physical objects and were presented in exact scale.

Each test trial was constructed from the names of two objects in the layout and required participants to point to an object (e.g., “Imagine you are facing the cell phone; please point to the ball”). The first object established the imagined heading (e.g., cell phone) and the second object was the target (e.g., ball). Participants pointed with a tracked hand-held wand.



Figure 5.1. The eight virtual objects used in the experiments.



Figure 5.2. Layout of objects used in the experiments. During the learning phase, half of the participants faced the cell phone and the other half faced the notebook.

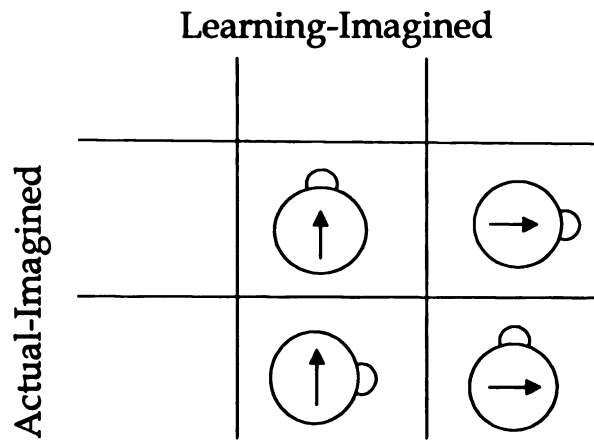


Figure 5.3. Design of experiments: Head-nose icons indicate actual headings; arrows indicate imagined headings. Headings and differences between them are measured counter-clockwise to maintain consistency with previous experiments.

The design is illustrated in Figure 5.3. The independent variables are (a) the angular difference between the learning heading and the imagined heading at the time of test and (b) the angular difference between the actual body heading and the imagined heading at the time of test. As is shown in Figure 5.2, to factorially manipulate these two variables, participants had two actual body headings at the time of test: One was the same as the learning heading (e.g., actually facing the cell phone), and the other was 90° different from the learning heading (e.g., actually facing the book). At each actual heading, participants had two imagined headings: One was the same as the learning view (e.g., “Imagine you are facing the cell phone”), and the other was 90° from the learning view (e.g., “Imagine you are facing the book”). Hence, as illustrated in Figure 5.2, the actual body heading was the same as the learning heading when the distances of the learning–imagined and the actual–imagined were the same (either both were 0° or both

were 90°), or it was 90° different from the learning heading when the distances of the learning–imagined and the actual–imagined were different (one was 0° and the other was 90°). Both of these variables were manipulated within participants. At each actual body heading, participants had 14 trials (pointing to each of the seven objects, except the imagined facing object at each imagined heading) in a random order. Participants would imagine themselves or the scene rotating 90° when the imagined heading was 90° from their actual heading (e.g., they believed they were actually facing the cell phone but were required to imagine facing the book). According to a study by Wraga et al. (2000), most people would rotate their body. However, participants were not explicitly instructed to adopt body or scene rotation when the imagined heading was different from the actual heading because that is beyond the scope of this study and would not change the results.

During the learning phase, half of the participants were randomly assigned to face the cell phone and the other half faced the book. This design counterbalanced the pointing direction across all four conditions (as is illustrated in Figure 5.3) and ensured that all conditions were equally difficult in terms of the pointing response. The order of the actual body headings at test time was also counterbalanced across participants: Half of them kept their learning orientation in the first block of pointing and then turned 90° for the second block; the other half performed in the reverse order. The primary dependent variables were pointing latency and pointing accuracy. Pointing directions were calculated in terms of the participants' facing direction.

5.2.1.3 Procedure

Participants were randomly assigned to each body-heading combination at test time, with the constraint that each group contained an equal number of men and women.

After providing informed consent, participants were trained in how to point from the imagined heading, which is either the same as or different from their actual heading. After participants understood how to conduct the pointing judgment, the experimenter escorted them to the learning room. To remove any potential orientation influence due to environmental structures, which may be represented in spatial memory, participants were blindfolded while being escorted into the learning room and to the learning position.

When the participants were standing in the learning position and facing the learning direction, the blindfold was removed. Then the participants were instrumented with the AR hardware system. The experimenter put a binder with a tracker on the participants' waist, placed the HMD with a tracker on their head, and handed them a pointing wand with a tracker. They were instructed to press a button on the wand when they felt they were pointing to the target object accurately. At this point, the learning phase began. Participants were instructed via earphones to point to all objects twice in a row with visual guidance (e.g., "Please point to the ball") to get used to the wand. Participants used earphones throughout the experiment to avoid any spatial references resulting from sound source location. After that, they were allowed to study the layout for 30 seconds, and then were asked to keep their eyes closed and to point to the objects named by the system. Participants performed five study–test sequences and were able to point to all of the objects accurately (within 15°). The audio cues were prerecorded so as to ensure consistency between subjects.

After participants had learned the layout, they were blindfolded and adopted the first actual body heading. Participants always stood at their learning position but turned their body if the actual body heading was different from the learning view. Test trials

were presented and participants were asked to point with the wand as accurately as possible before they pressed the button. The tracker on the wand recorded the pointing direction; pointing latency was recorded from the onset of the target object cue to the button-press. After they finished all 14 trials, they adopted the second actual body heading (turned by the experimenter) and repeated the same 14 trials.

5.2.1.4 Results and Discussion

Pointing accuracy and pointing latency as a function of actual–imagined distance and learning–imagined distance are presented in Table 5.1. The means for each participant and for each condition were analyzed through the use of repeated-measures analyses of variance (ANOVA) in terms of actual–imagined distance (0° and 90°) and learning–imagined heading (0° and 90°).

	A-I = 0°				A-I = 90°			
	Latency		Accuracy		Latency		Accuracy	
L-I	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
0°	3.821	2.139	14	8	4.515	2.081	15	9
90°	4.384	1.683	15	11	5.117	2.097	17	11

Table 5.1. Pointing latency (in seconds) and pointing accuracy (in degrees) as a function of Actual-Imagined (A-I) distance and Learning-Imagined (L-I) distance in Experiment 1.

	F(1,15)		Cohen's <i>f</i>	
Source	Latency	Accuracy	Latency	Accuracy
L-I	4.52*	.32	.55	.15
Error	1.20	123.60		
A-I	11.11**	.77	.86	.23
Error	.73	39.11		
L-I x A-I	.01	.19	.00	.11
Error	1.33	39.75		

* $P < .05$, ** $p < .01$

Table 5.2. Analysis of variance results for pointing latency and pointing accuracy in Actual-Imagined (A-I) and Learning-Imagined (L-I) conditions in Experiment 1.

The ANOVA results for pointing accuracy and pointing latency are presented in Table 5.2. In angular error, no main effect was significant. People were highly accurate in all conditions. In pointing latency, both main effects of learning–imagined and actual–imagined were significant, whereas the interaction between them was not.

The most important result of Experiment 1 was that pointing latency was shorter when the actual and the imagined headings were the same (0°) than when they were different (90°). This result indicates that people cognitively update the location of the virtual object when they rotate their body. In other words, humans use an environment-stabilized reference frame to access information arrays. The evidence for this is the cost in latency that was incurred by the need to align the egocentric front with the facing

object specified in the pointing judgment. Given that the imagined heading is the same, the pointing latency should be the same when they are facing the learning view as when they are turned 90° from it when participants use a body-stabilized reference frame. This result showed that turning 90° from the learning view at test time benefited the imagined heading of 90° but had the reverse affect on the imagined heading of 0°.

The second important finding was that pointing latency was shorter when the imagined and learned headings were the same (0°) than when they were different (90°). This result indicates that people represent the location of the virtual object with a reference frame selected by the learning view; that is, spatial memory is orientation dependent.

Both of these results were consistent with the research of spatial updating of physical objects (Mou *et al.* 2004b), suggesting that the spatial cognition system codes and processes spatial locations of virtual objects presented in AR environments using the same coding and processing as for physical objects.

5.3 Experiment 2: Adaptation of Egocentric Frame with Prior Experience

The results of Experiment 1 indicate that participants used an environment-stabilized frame of reference to access the location of virtual objects if they had never experienced the possibility that objects can also be attached to the body (egocentric reference frame) in virtual and AR environments. In Experiment 2, we examined whether direct experience of virtual objects attached to an egocentric reference frame in which objects translate and rotate with the moving body (a condition rarely experienced in the physical world) would stop participants from updating their actual heading with respect to the layout but would, instead, cause them to use a body-stabilized reference frame for

other layouts. Evidence of this effect would suggest that users are capable of learning to use and update arrays of menus and objects organized around their moving body and that an egocentric infospace would be accepted and processed cognitively in such a way that it would be an efficient information presentation medium.

5.3.1 Methodology

5.3.1.1 Participants

Participants were 8 female and 8 male undergraduates at Michigan State University who participated voluntarily as partial fulfillment of course requirements.

5.3.1.2 Materials, design, and procedure

The materials, design, and procedure of Experiment 2 were similar to those of Experiment 1 except a training session was added before participants learned the experimental layout of eight objects.

During the training session, five virtual objects were presented. Participants were instructed to look at the locations of all objects. After they saw all of them, they were asked to turn left and look at the locations from the new viewing direction. The objects were simultaneously rotated in space so as to maintain their position and orientation relative to the participant's body. The subjects then turned back to adopt the original orientation and took a look at the locations of the objects. The process was repeated with a right turn, again maintaining object position and orientation relative to the subject's body. Finally, they were instructed to return to the initial orientation. The training session lasted approximately 2 minutes. The experimenter did not comment on or verbally explain the behavior of the virtual objects. Learning about the object behavior was through observation only.

Following the training session, the learning session started. The learning session was identical to that of Experiment 1.

5.3.2 Results and Discussion

Pointing accuracy and pointing latency as a function of actual–imagined distance and learning–imagined distance are presented in Table 5.3. The means for each participant and each condition were analyzed in repeated-measures ANOVAs in terms of actual–imagined distance (0° and 90°) and learning–imagined heading (0° and 90°). The ANOVA results for pointing accuracy and pointing latency are presented in Table 5.4. In angular error, no effect was significant. People were highly accurate in all conditions. In pointing latency, only the main effect of learned–imagined was significant.

	A-I = 0°				A-I = 90°			
	Latency		Accuracy		Latency		Accuracy	
L-I	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
0°	4.077	2.030	13	9	4.510	2.947	19	17
90°	5.513	3.347	17	7	5.777	3.441	19	12

Table 5.3. Pointing latency (in seconds) and pointing accuracy (in degrees) as a function of Actual-Imagined (A-I) distance and Learning-Imagined (L-I) distance in Experiment 2.

	F(1,15)		Cohen's <i>f</i>	
Source	Latency	Accuracy	Latency	Accuracy
L-I	20.21**	.26	1.16	.13
Error	1.45	77.96		
A-I	2.40	3.48	.40	.48
Error	.81	73.91		
L-I x A-I	.05	1.07	.05	.27
Error	2.50	56.05		

** $p < .01$

Table 5.4. Analysis of variance results for pointing latency and pointing accuracy in Actual-Imagined (A-I) and Learning-Imagined (L-I) conditions in Experiment 2.

The most important finding of Experiment 2 was that the effect of the angular distance between the imagined heading and the actual heading on pointing latency was not significant. Although failing to reject the null hypothesis is not the same as demonstrating the validity of the null hypothesis, it is safe to conclude *that the effect of the actual–imagined heading on pointing latency decreased after people had a brief exposure to a body-stabilized display*. The difference in pointing latency between actual–imagined (0° and 90°) decreased from 713 ms in Experiment 1 to 394 ms in Experiment 2. The effect size *f* on pointing latency consistently decreased from .86 in Experiment 1 to .40 in Experiment 2. It is hard to exclude the possibility that some participants showed the actual–imagined effect and others did not because this is not an individual-based

experiment. In general, however, the results indicate that participants were able to use the body stabilized, egocentric reference frame to access information for the location of an array of virtual objects after only 2 min of prior exposure to the location of an array of virtual objects.

5.4 Experiment 3: Adaptation to an Egocentric Frame with Oral Instruction

In Experiment 3, we examined whether participants who were instructed that the layout was stabilized with respect to their body would stop updating their actual heading with respect to the layout and would, instead, adopt a body-stabilized, egocentric reference frame.

5.4.1 Methodology

5.4.1.1 Participants

Participants were 8 female and 8 male undergraduates at Michigan State University who participated voluntarily as partial fulfillment of course requirements.

5.4.1.2 Materials, design, and procedure

The materials, design and procedure were similar to Experiment 1 except for the following two modifications:

1. Prior to the physical turn of the participants during the testing phase, they were given a body-stabilized instruction (e.g., “When you physically turn your body, the objects on the floor will move as you turn. Hence, after you turn right, you will be still facing the cell phone”).
2. A new motion tracking system, InterSense IS-900, was used due to an upgrade to the experiment facility. The new tracking system performed identically to the original system with the exceptions of a considerably increased range and

slightly decreased latency, and, thus, is not likely a significant factor in these experiments.

5.4.1.3 Results and Discussion

Pointing accuracy and pointing latency as a function of actual–imagined distance and learning–imagined distance are presented in Table 5.5. The means for each participant and each condition were analyzed in repeated-measures ANOVAs in terms of actual–imagined distance (0° and 90°) and learning–imagined heading (0° and 90°). The ANOVA results for pointing accuracy and pointing latency are presented in Table 5.6. In both angular error and pointing latency, only the main effect of learned–imagined was significant. The results clearly indicate that *after being instructed that the objects were arrayed around the body in a body-stabilized display, people used a body-stabilized frame of reference to access the information array.*

	A-I = 0°				A-I = 90°			
	Latency		Accuracy		Latency		Accuracy	
L-I	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
0°	4.077	2.030	13	9	4.510	2.947	19	17
90°	5.513	3.347	17	7	5.777	3.441	19	12

Table 5.5. Pointing latency (in seconds) and pointing accuracy (in degrees) as a function of Actual-Imagined (A-I) distance and Learning-Imagined (L-I) distance in Experiment 3.

	F(1,15)		Cohen's <i>f</i>	
Source	Latency	Accuracy	Latency	Accuracy
L-I	18.68**	9.98**	1.12	.82
Error	4.72	327.05		
A-I	.09	2.43	.08	.40
Error	1.51	89.90		
L-I x A-I	.08	1.38	.07	.30
Error	4.62	133.95		

** $p < .01$

Table 5.6. Analysis of variance results for pointing latency and pointing accuracy in Actual-Imagined (A-I) and Learning-Imagined (L-I) conditions in Experiment 3.

5.5 Discussion

Current 3D graphics and tracking technology allow designers to display information arrays around a mobile AR user with respect to a body-stabilized or an environment-stabilized frame of reference. There have been no prior studies conducted to investigate which reference frame mobile users use and what factors may influence choices of reference frame. This study, through the use of the paradigm developed to investigate human spatial memory and spatial updating in physical environments (Mou *et al.* 2004b), suggests that users with no prior experience of mobile AR systems tend to use an environment stabilized reference frame to access information arrays presented in AR environments. In other words, people expect the information arrays of virtual objects in

AR environments to behave like arrays of objects in physical environments (i.e., when they rotate their body, objects stay in their locations relative to the physical environment). This study also suggests that users who briefly experience the egocentrically centered display of virtual objects or those who are instructed that the display is egocentrically centered are capable of *quickly adopting a body-stabilized reference frame to code and access the locations of virtual objects in the physical environment*.

Why do naïve users think the locations of the virtual objects are stabilized with respect to the environment? One apparent explanation is that from birth, human beings perceive that the locations of objects in the environment as independent of their own locomotion and, thus, the relationship between their body's locomotion and changes of self-to-object relations are represented in their cognitive system. To efficiently locomote in an environment where objects are not always visible, humans have to develop the ability to update locations of objects in the environment without visual guidance (Rieser *et al.* 1986; Rieser 1989; Presson *et al.* 1994; Farrell *et al.* 1998; Simons *et al.* 1998; Wang *et al.* 1999; Sholl *et al.* 2002; Waller *et al.* 2002; Mou *et al.* 2004b). People couple their motions and locomotion with an automatic spatial updating of the representation of object locations. They do so by coupling their locomotion with the perception of change in the spatial relations between the body and objects in the environment during their interaction with the environment (Rieser, Pick and Ashmead 1995; Rieser 1999). People with no prior experience in mobile AR systems simply interpret the relation between their locomotion and the locations of virtual objects with the mental model they use to interpret the physical world.

On the other hand, the results of Experiments 2 and 3 demonstrate that this lifetime experience with physical objects can be quickly replaced with a model of virtual object arrays that move with the body. In Experiment 2, participants perceived that the locations of virtual objects stayed stationary with respect to their body rotation for only 2 minutes. Their spatial updating behavior indicated that people in general tend to use body-stabilized reference frame to code and access the locations of virtual objects after experiencing the behavior of these objects in the new AR layout. This implies that people couple their motions and locomotion with a cancellation of the spatial updating of the representation of object locations. They do so during their interaction with the environment by coupling their locomotion with the perception of “unchanged” in the spatial relations between the body and objects in the environment.

The quickness with which participants adapted to the egocentric array of object locations is very promising, as far as use of that Peripersonal Infospace to hold digital information in AR environments is concerned. It was speculated that people might have a mental model in favor of a body-stabilized reference frame that can accommodate arrays of virtual objects that move with the body even though they have no visible means of attachment to the body. This consideration was supported by the results of Experiment 3, which showed that even without any direct experience, and with only oral instruction that the objects were fixed relative to the body (body-stabilized frame of reference), people were able to use body-stabilized frames of reference to code and access the locations of virtual objects. The results of both Experiments 2 and 3 also suggest the spatial cognitive system is highly flexible with respect to spatial updating.

Can users of AR systems remember and make use of arrays of three-dimensional objects that move around the body even when they are more than 1 m away from the body? The results of these studies suggest that high quality, mobile AR interfaces may be able to leverage the capacity of human spatial memory and spatial updating mechanisms for efficient access to information items around the body. In this study, we attempted to (a) identify the default frame of reference in coding virtual objects in a high-quality AR mobile system and (b) determine whether experience and oral instruction could alter it. Further studies should investigate how people encode the locations of virtual objects on occasions in which both body-stabilized and environment-stabilized frames of reference are necessary. It remains to be seen whether the updating of these virtual objects interferes with the updating process for objects in the physical environment. This notwithstanding, the current study provides answers to the questions raised in the introduction: Users with no prior experiences in mobile AR systems tend to use environment-stabilized reference frames to encode and access information arrays around their body. Evidently, experiences with or oral instructions of a body-stabilized display allow users to adopt a body-stabilized frame of reference instead.

6 Evaluation of Perceptual Asymmetric Effects in Egocentric Infospaces

Perceptual asymmetric effect can potentially impact the human cognition system in various ways, such as in reaction time, perception, induced emotion, and semantic meaning. While perceptual asymmetries are well-known effects in cognitive psychology and can be easily demonstrated in laboratory settings, it is not clear at all that these effects can be directly utilized in user interface scenarios. Empirical study on perceptual asymmetry effects *in an application setting* is needed before applying them to the design of interfaces. Two experiments were conducted to investigate the practical impact perceptual asymmetric effects on actual tasks.

6.1 Experiment 4: Evaluation of Left vs. Right Instruction Presentation

An experiment was conducted to evaluate asymmetrical effects of graphical and text instructions placed on the left or right side of a head-stabilized reference frame with an emphasis on the impact on task completion time. According to research in psychology, the left visual field is superior for word recognition (Melville 1957; Bauma 1973; Axelrod, Haryadi and Leiber 1977; Young and Ellis 1985; Ellis, Young and Anderson 1988) and language processing (Sperry 1961; Gazzaniga and Sperry 1965), while the right visual field is superior for geometric patterns and visual orientation matching (Atkinson and Egeth 1973). Since visual stimulus presented on one side of the head-stabilized reference frame will predominantly fall on the same side of the visual field. it is predicted that:

H1: Task completion time for graphical instruction presented on the right side of the head-stabilized reference frame will be significantly shorter than on the left side, and

H2: Task completion time for text instruction presented on the left side of the head-stabilized reference frame will be significantly shorter than on the right side.

6.1.1 Methodology

A within-subjects experiment was designed. There were two dependent variables: position of the instruction on the head-stabilized reference frame (left vs. right), and type of instruction (graphic vs. text). The independent variable is the time of completion of the experimental task. Four experimental conditions were created: (1) Graphic instructions presented to the left, (2) Graphic instructions presented to the right, (3) Text instructions presented to the left, and (4) Text instructions presented to the right

6.1.1.1 Stimulus Materials

Participants were asked to complete a task of arranging Duplo blocks into a spatial pattern and then pressing a button of a specific color according to instructions presented in a head-stabilized reference frame. Arranging Duplo blocks into spatial patterns was chosen as a task in the experiment to minimize bias towards a population with expertise in a certain knowledge related to a task. In each trial, participants were asked to acquire 5 to 15 Duplo blocks of different colors from an unsorted bin and arrange them into the patterned presented in the instruction. Figure 6.1 shows examples of instruction presented and the completed task.

(a)

Row 1: Yellow, Yellow, Blue, Red, Blue,
Yellow, Yellow
Row 2: Yellow, Blue, Blue, Yellow, Blue
Button: Blue

(c)

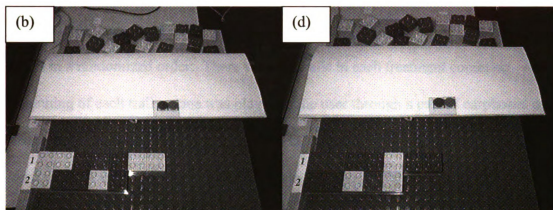


Figure 6.1. Examples of instruction and the completed task. Example of text instruction is shown in (a) instruction and the completed task is shown in (b), and example of graphic instruction is shown in (c) and the completed task is shown in (d).

6.1.1.2 Participants

Participants were 8 undergraduate students at Michigan State University who participated voluntarily as partial fulfillment of course requirements. None of them had previous experience in any AR environment.

6.1.1.3 Experimental Equipments

Visual cues were displayed in stereo with the Sony Glasstron LDI-100B head-mounted display, and audio stimulus materials were presented using a pair earphones.

6.1.1.4 Procedure

Participants were first introduced to the experimental procedure and equipments, and then entered the pretest environment. A few example instructions were presented to the participants and the experimenter explained the tasks in the experiment. When participants indicated that they understood the experimental procedure and the task, the experiment began where they experienced each interface treatment condition (graphical instruction of left, graphical instruction of right, text instruction of left, text instruction of right) in a randomized order. There were 12 trials in each treatment condition. At the beginning of each trial, a tone was played to the user through a pair of earphones and the visual instruction was displayed on the HMD according to the treatment condition. Participants were to arrange the Duplo blocks into the spatial pattern according to the instruction displayed on the HMD and then press the button of the color as specified on the instruction.

6.1.1.5 Measurements

Task completion time in milliseconds was measured as the time it took for participants to press the button following the onset of an audio cue tone.

6.1.2 Results

The mean and standard deviation of task completion time for each condition is summarized in Table 6.1. A general linear model repeated measure analysis was conducted to test the effect of stimulus position on task completion time. There was no statistical significant effect of stimulus position on task completion time for both graphical and text instruction, $F(1, 8) = 3.213$, $p = 0.116$ for graphical instruction; $F(1, 8) = 0.930$, $p = 0.372$ for text instruction.

	Graphical Instruction on Left Side	Graphical Instruction on Right Side	Text Instruction on Left Side	Text Instruction on Right Side
Mean	20308 ms	22190ms	22709ms	21122ms
S.D.	8316	9843	10068	7788

Table 6.1. Task completion time and standard deviation in Experiment 4.

The descriptive statistic indicates an advantage for graphic instruction placed on the left side and text instruction placed on the right side. However, none of these results has any statistical significance.

6.1.3 Discussion

In contrary to the hypothetical prediction, graphic instruction presented on the left side and text instruction presented on the right side had a shorter task completion time in general. However, the experimental results did not achieve statistical significance. One explanation of this effect is that effects of perceptual asymmetries apply to the visual field of the retinal image only.

Even though visual stimuli placed on one side of the head-stabilized display are predominantly fall on the same side of the visual field, the retinal image may move to another side due to eye movement. Hence, simply placing information items to one side does not guarantee the image will be projected to that side of the visual field exclusively. Furthermore, the reaction time for perceptual effects of bilateral asymmetry is measured in milliseconds. The perceptual advantage of bilateral asymmetry could be relatively

insignificant when compared to other cognitive and psychomotor process such as sorting and motor action planning and execution. A perceptual advantage measured in milliseconds does not have a significant impact on a task that spans 5 to 10 seconds. In conclusion, Perceptual asymmetric properties in reaction time are not robust, and too subtle to have practical effects on reaction to stimuli in AR and other information displays. These effects can only be used sparsely for information placement in egocentric infospaces.

6.2 Experiment 6: Emotion and Semantic Meaning

Semantic meaning is likely to be mapped to proximity. An experiment was conducted (Biocca, Lamas, Gai, Brady and Tang 2001; Biocca, David, Tang and Lim 2004) to explore how semantic meaning of virtual objects and agents change with the location around the body. Do positions in space around the body carry meaning? Is the spatial location of an object part of its connotative meaning?

6.2.1 Related Works

Approach and avoidance fields around animals and humans are well documented. The work of proxemics is well known in communication research. The term "proxemics" was coined by Edward Hall in 1963 (Hall 1963; Hall 1966) when he investigated human's use of personal space in the communication and social context. His theory suggests that human maintain different levels of distance to different people, agents and objects in space. For example, people in the United States maintain a 6 to 18 inches distance as an intimate distance for embracing, 1.5 to 4 feet as a personal distance for good friends, and a 4 to 12 feet distance as social distance for everyday conversation. A violation of these *comfort distances* (e.g. a stranger stepping into the 1.5' – 4' personal distance) could

influence the perception of the intention of other people, and could trigger different emotional response and semantic meaning. This research suggests that spatial location of agents, people and objects in space relative to the body will have meaning, especially as the location crosses a threshold into the *private space* around the body.

Looking at the semantic oppositions within language, there is good support for this spatial semantic asymmetry. Cultures worldwide map meaning to the high-low dimension of space. For example, within most Greco-Roman languages, it is very consistent that “up” or “above” is perceived with positive meanings and “down” and “below” with negative. There is some linguistic evidence for the semantic asymmetries of spatial location.

There is also some evidence in neuroscience for differences in the processing of peripersonal space and extrapersonal space. Neurophysiological studies in brain-injured patients have shown that lesions in different brain regions can lead to asymmetrical neglect for near or far spaces consistent with peripersonal and extrapersonal space (Berti, Smania and Allport 2001).

Humans also spontaneously respond to affordances in the environment that are correlated with sentient beings such as other humans and animals (Sheehan and Sosna 1991). Mediated embodiments such as pictures, computer characters, moving robots, and other representations of “apparently sentient” others can automatically trigger social presence responses (Reeves and Nass 1996). The philosophical and psychological concept of agency has many subtle dimensions (McCann 1998; Bratman 1999). The concept of agency, defined as the state of being in action or of exerting power, is central to the issue of the volitional or intentional force that drive the actions of an entity. This

property of potentially acting within a space may make the spatial location of agents more salient, and therefore more meaningful, to the user of a virtual environment.

6.2.2 Methodology

A 5 by 2 by 2 within-subjects experiment was designed with three within-subjects factors: (1) location around the body, (2) distance from the body, and (3) type of object. Location around the body had five levels defined by spatial location. Distance from the body had two levels, near and far (see Figure 6.2). And finally, the factor, agency, had two levels, *agent representation* (a 3D anthropomorphic head) and *object representation* (i.e., a simple golden sphere) (see Figure 6.3).

6.2.2.1 Participants

13 undergraduate students participated in the study voluntarily for class credit. All participants were right handed.

6.2.2.2 Stimulus Materials

Two types of stimulus were used to manipulate the perception of agency, a 3D human head or a golden sphere, as illustrated in Figure 6.3. In each experimental trial, either the sphere or the head appeared in one of the ten predefined spatial locations. The ten predefined spatial location, five for near space and five for far space, are as shown in Figure 6.2. The left, right, up and down positions are 30° deviated from the center location. The near locations are measured 3 feet from the participant's body, and the far locations are measured 10 feet from the participant's body.

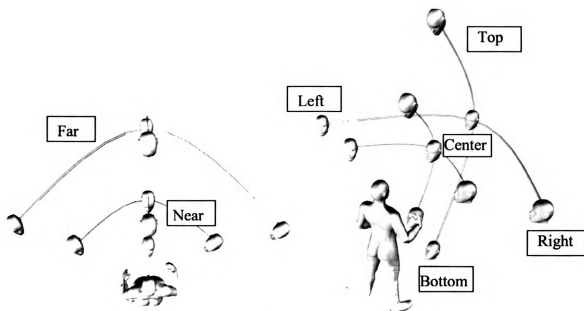


Figure 6.2. Ten predefined locations around the body. The five locations in the near space are 3' away from the body. The five locations in the far space are 10' from the body. The above, below, left and right locations is deviated 30° from the center location.

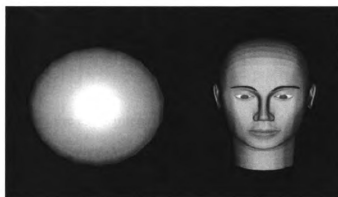


Figure 6.3. The two stimulus material used in the experiment. The golden sphere used for object representation is shown on the left side. The human head used for agent representation is shown on the right side.

Head motion was tracked with a Polhemus Fastrak magnetic tracker. Stereo graphics were rendered in real time on the basis of the data from the tracker using a Research V8 stereoscopic HMD. An SGI™ Onyx® Reality Engine² running with two graphics pipes and MultiGen® SmartScene™ software was used to render the stimulus materials in real time.

6.2.2.3 Measurement

A set of semantic meanings was measured using four bipolar measurements from the classic semantic differential instrument (Osgood, Suci and Tannenbaum 1957).

These four items were selected based on the result of a pilot test. Those that were most sensitive to variations in the spatial location of objects were retained. The four bipolar measurements selected are superior-inferior and relevant-irrelevant in evaluative factor, (2) urgent-not urgent in potency factor, and aggressive-peaceful in activity factor. Each measurement item provided the anchor points on a seven point scale. For example, the superior-inferior measurement asked the participants to answer the following question: “How superior or inferior is the object? Very superior, moderately superior, slightly superior, neutral, slightly inferior, moderately inferior, or very inferior.”

6.2.2.4 Procedure

Participants entered the experiment room and were briefed about the equipments and were assisted to mount the HMD. At the beginning of each trial, one of the ten predefined locations was randomly selected, and either a head or a sphere appears in that location. Participants then observed the object for about 10 seconds. A questionnaire containing the 4 items semantic measurement then appeared in front of the participants’ visual field. The subjects read the questionnaire and respond orally to the experimenter.

After the 4 questions were completed, the trial ended and next trial began. There were twenty trials total.

6.2.3 *Results and Analysis*

Table 6.2 summarizes the mean of the four semantic differential measurements for the 3 experimental factors. The data analysis was conducted in four steps. In each step, a different semantic-differential measure was entered as the dependent variable and the data analyzed using a 5 (Position: center, left, right, top, bottom) x 2 (Distance: close, far) x 2 (Agency: agent, ball) within-subject repeated measures analysis of variance. Due to the complexity of the design, each of the four measures was treated separately so that higher-order interactions might be more interpretable. Correlation analysis of the measures revealed that the four items were significantly correlated with pair-wise correlations ranging from .55 to .75. However, a composite score of the four items was not examined because the purpose of this study was to examine different aspects of the semantic space, rather than one overall evaluation.

	Overall	Distance		Position					Object	
		Near	Far	Center	Left	Right	High	Low	Face	Sphere
Superior (1)/ Inferior (7)	4.2	3.7	4.7	4.5	4.5	4.1	4.0	3.8	3.6	4.8
Relevant (1) Irrelevant (7)	4.0	3.7	4.5	4.3	4.4	4.0	3.6	4.2	3.4	4.7
Urgent (1) Not-Urgent	4.6	4.0	5.2	4.7	4.7	4.7	4.3	4.5	4.2	5.0
Aggressive (1)/ Peaceful (7)	4.2	3.8	4.5	4.3	4.3	4.2	3.8	4.2	3.6	4.7

Table 6.2. Means for the Different levels of the 3 Experimental Factors

In within-subjects designs it is highly likely that the sphericity assumption is violated, as a result inflating the degrees of freedom. To adjust for this, the Huynh-Feldt correction was applied to the degrees of freedom, which is the reason that some of degrees of freedom presented in this section have decimal values. For each of the univariate tests, the corresponding multivariate test was also examined. The univariate and multivariate tests were nearly identical in most cases, with two exceptions. In one case, the multivariate test was significant, but the univariate test was slightly above the .05 level, whereas in the other instance, the univariate test was significant, but multivariate test was above the .05 level.

6.2.3.1 Superior/Inferior

When the ratings of the semantic differential superior/inferior were analysed, the main effects for Position, $F(4, 48) = 2.65, p = .04$, Distance, $F(1, 12) = 28.27, p = .000$, and Agency, $F(1, 12) = 17.54, p = .04$, were found to be statistically significant. Nearer items that appeared in the near space were rated as superior ($M = 3.7, SD = 1.8$), in comparison to farther items that appeared in the far space ($M = 4.7, SD = 1.7$). Also, the 3D human face was rated as superior ($M = 3.6, SD = 1.7$) in comparison to the golden sphere ($M = 4.8, SD = 1.7$). To examine the difference between levels that contributed to the main effect for Position, two within-subjects contrasts were used. The contrast comparing the left vs. right positions was tending toward significance, $F(1, 12) = 3.70, p = .07$; whereas the contrast for top vs. bottom was not significant. Items that appeared in the right field ($M = 4.2, SD = 1.9$) had higher superiority ratings compared to items that items presented in the left field ($M = 4.5, SD = 1.8$), offering some evidence of a left-right asymmetry.

6.2.3.2 Relevant/Irrelevant

The relevance of the object in space was analysed. The main effects for Position, Distance and Agency were statistically significant, $F(3.3, 39.4) = 3.36, p = .03$ for Position; $F(1, 12) = 17.49, p = .001$ for Distance; and $F(1, 12) = 19.53, p = .001$ for Agency. Among the interactions, only the two-way interaction for Position x Distance was significant, $F(3.3, 39.3) = 2.91, p = .042$. An interaction contrast revealed that Position x Distance interaction was accounted for by the differences between the left vs. right positions, $F(1, 12) = 8.09, p = .015$. Nearer items within the peripersonal space ($M = 3.7, SD = 1.6$), were rated as being more relevant than items outside the peripersonal

space ($M = 4.5$, $SD = 1.7$). By the same token, the 3D face model ($M = 3.4$ $SD = 1.6$) was rated as more relevant than the golden sphere ($M = 4.7$, $SD = 1.5$). While the left-right asymmetry observed in the superiority ratings was also apparent for relevance, a two-way interaction between distance and left-right positions was further analysed. Figure 6.4 illustrates this two-way interaction, that items within peripersonal space were rated more relevant than items outside peripersonal space, the left-right asymmetry was observed only for the items within the peripersonal space, but not for items outside the peripersonal space.

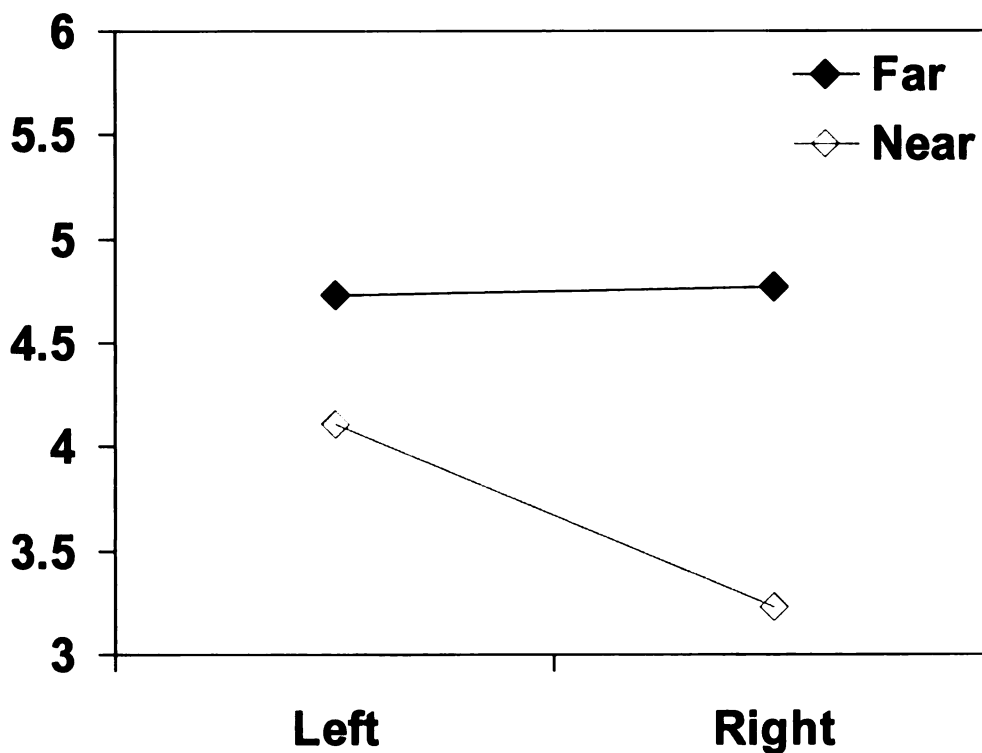


Figure 6.4. Relevant-Irrelevant by distance and position of objects

6.2.3.3 Urgent/Not Urgent

Urgency has a statistically significant effect on Distance and Agency, $F(1, 12) = 42.55$, $p = .000$ for Distance and $F(1, 12) = 7.86$, $p = .016$ for Agency. In addition, a Position x Distance two-way interaction was statistically significant, $F(4, 48) = 2.58$, $p = .049$. The Position x Agency two-way interaction was trended toward significance, $F(2.8, 34) = 2.63$, $p = .069$. Nearer items ($M = 4.0$, $SD = 1.7$) were perceived as more urgent than items outside peripersonal space ($M = 5.2$, $SD = 1.4$). The effect of distance was asymmetrical between the left and right fields, with the effect of distance being more pronounced in the right field than in the left field.

Figure 6.5 illustrates this asymmetric effect. Also, the 3D face ($M = 4.2$, $SD = 1.7$) evoked more urgency than the golden sphere ($M = 5.0$, $SD = 1.6$) and the effect of agency was more pronounced when the objects were located away from the center. At the center, there was no difference in the perceived urgency for the sphere and 3D face, however, once when these objects appeared in the right or left fields, a difference trending toward significance emerges ($p = .07$), with the face being perceived as more urgent than the sphere. Figure 6.6 illustrates this effect of Urgency by position.

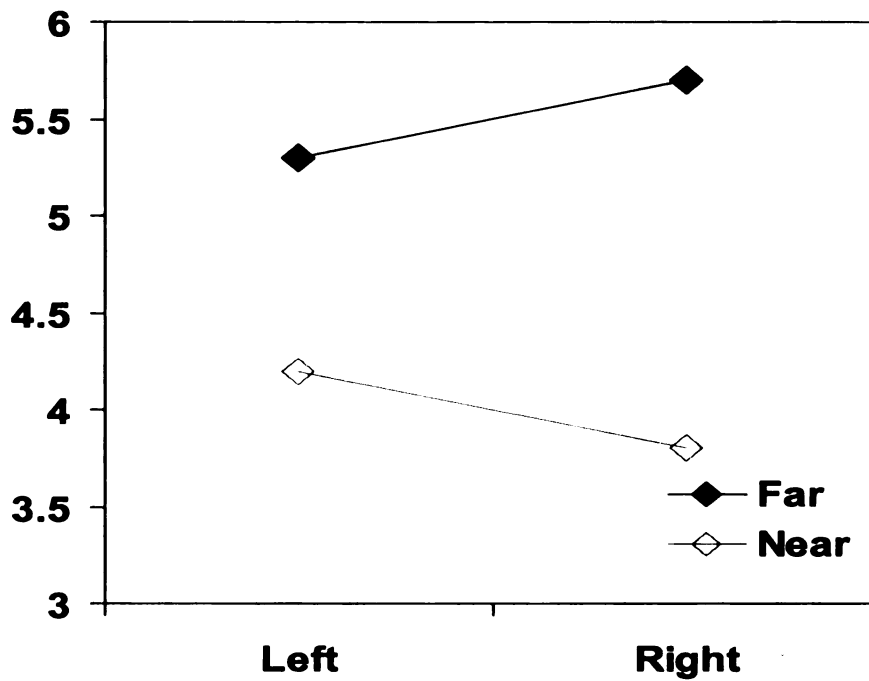


Figure 6.5. Urgent-Not urgent by distance and position.

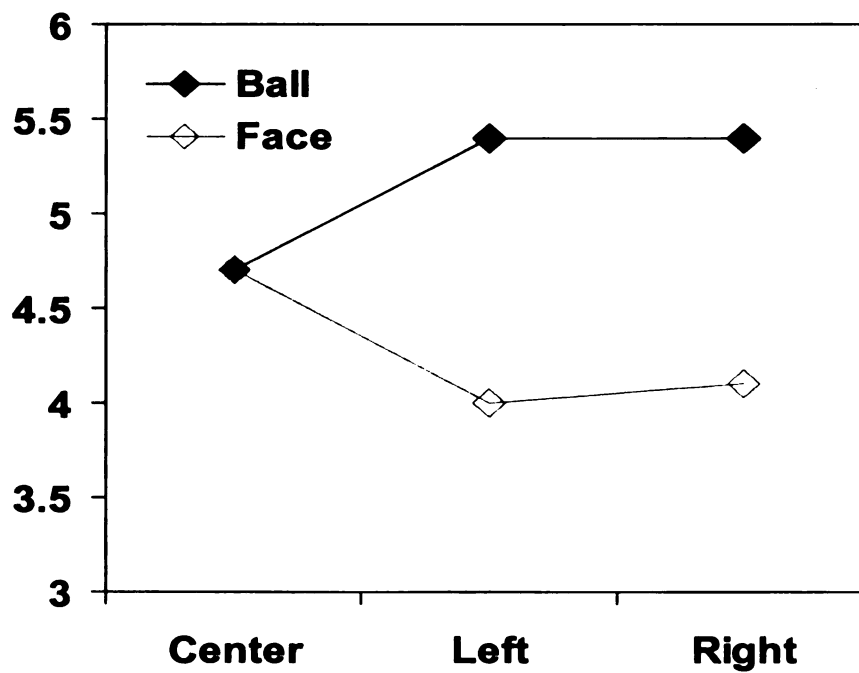


Figure 6.6. Urgent-Not Urgent by object and position.

6.2.3.4 Aggressive/Peaceful

When ratings of aggressive/peaceful were analysed, the main effects for distance and agency were found statistically significant, $F(1, 12) = 7.14$ for distance and $p = .02$, $F(1, 12) = 9.73$, $p = .009$ for agency. Furthermore, objects in near space ($M = 3.8$, $SD = 1.8$) were seen as more aggressive than objects in far space ($M = 4.5$, $SD = 1.7$). Agency also had a sizable impact on the aggressive/peaceful rating, with the 3D face ($M = 3.6$, $SD = 1.7$) rated more aggressive than the golden sphere ($M = 4.7$, $SD = 1.7$).

6.2.4 Discussion

The clearest finding from this study is that location in virtual space appears to have semantic meaning. Participants ascribed differences in meaning, as measured by items from all the dimensions typically captured by the classic semantic differential scale (Osgood *et al.* 1957), as objects and agents changed position in space. Depending on their location, agents and objects differed in their superiority and relevance (evaluative factor), urgency (potency factor) and in their level of aggression-peacefulness (activity factor). It is relevant to note that Osgood's semantic differential was conceptualized using a spatial metaphor of semantic spaces. In this study it appears that participants are actually mapping semantic properties and connotations to locations in space.

6.2.4.1 The Effect of Distance in Connotative Semantics

One of the strongest effects observed in this study was the shifts in connotative meanings when an object is located in the near space. *Objects in the near space appeared to be more relevant, superior, urgent, and aggressive.*

Some of these findings might be predicted with the literature on proxemics (Hall 1963). But that literature does not explain why objects, as opposed to people, should also

have a shift in connotative semantic properties within peripersonal space. One possibility is that the sphere, which floated in space, might have been seen as an agent, or at last closer in agency to a three-dimensional head. But this is not a plausible explanation, as it would suggest that the two would have been seen as largely equivalent. The three-dimensional head was seen as significantly different from a sphere on all our measures. Further research is required to examine this specific effect.

6.2.4.2 On the Semantic Volatility of Agent's Spatial Location

The simplest interpretation of the findings regarding agency is: People are more meaningful than objects. The virtual head was perceived as more superior, relevant, urgent, and aggressive than the virtual sphere. But the finding goes beyond this. The experiment did not use a person, but only a virtual head that did not display much in the manner of true agency or life like characteristics. Consistent with the work on virtual agents, the mere representation of agency, a simple and not very realistic three-dimensional virtual head, evokes more meaningful responses than the object, even when they are roughly equal in size and exactly equal in location. The users' perception of the ability of agents to act within the space may also evoke greater uncertainty and more volatile shifts in connotations. There may be uncertainty regarding about agents facing the user, but located on the side of the body. When agents moved from being directly in front of the body, to be on either side of the body, they are seen as more urgent.

6.2.4.3 Left-right asymmetries in the connotative semantics of virtual space

Some of the more robust findings in the literature are those that find left-right asymmetries in the visual processing of objects presented briefly within the visual field. This is typically interpreted as evidence for bilateral differences in information

processing and function across the brain's hemispheres. In this study, *objects falling on the left side of the body tended to have semantic properties with values that deviated from neutral relative to objects falling on the right side of the body*. This finding can be interpreted by properties of the right hemisphere. The right hemisphere is dominant for expressing and perceiving emotion, regardless of valence (Sackheim, Gur and Saucy 1978). Objects falling on the right side of the body tend to be perceived as neutral.

6.2.4.4 Beyond spatial location on the retina to location around the body

Most of those studies exploring differences in spatial location control location by placing objects within a specific location within the visual field and for a brief exposure. Objects fall on either the left or right hemiretina. They are not open to examination because they are presented for a very brief duration. The experiment put objects on the left and right of egocentric space. So any object could fall on either retina and processed by either hemisphere when observed directly. Nonetheless, the findings here are similar to visual field studies. One possible explanation for this effect is that the objects fall predominantly within one visual field, especially when on the eccentricities of the body. So they might be “predominantly” processed in one hemisphere or another. But effects argued from location within the visual field only would be very weak. It is also possible that mere location around the body regardless of where objects fall on the visual field has meaning. Such findings are suggested by work by Graziano *et al.* (Stein 1984; Graziano *et al.* 1995; Graziano and Gross 1998; Graziano 1999) with animals demonstrating the coding for egocentric location regardless of the location of the eye of observer. Location relative the body is coded in multimodal integration of space (Marks 1978; Stein 1984;

Marks and Armstrong 1994). These location may have some of the properties observed in visual field studies, and may not rely just on location on the retina.

6.3 Summary

Two experiments were conducted to evaluate the applicability of perceptual asymmetric effects for information display in Egocentric Infospaces. The experiment findings of the two experiments were mixed. Results of Experiment 4 suggest that reaction time based perceptual asymmetric effect is too subtle to have an effect on task temporal performance. Results of Experiment 5 suggest that spatial locations around the body are prescribed with semantic meanings. For example, connotations to a three-dimensional face presented in a virtual environment may be slightly different if this face is presented left or right, close or far in the virtual environment. AR interface designers may need to be aware that location may impart connotation and emotion to user's perception of the information objects, especially agents. The findings also suggest that spatial location may impart connotations to objects that do not necessarily have agency. But the shift in connotations of objects may be volatile and less influenced by spatial location than agents.

7 Directing Attention in Mobile AR Interface

One basic user interface functionality is the ability to direct attention to physical or virtual objects in the environment. Mobile, context-aware interfaces will often be tasked with directing attention to physical or virtual objects that are located anywhere in the environment around the user. Often the target of attention will be beyond the visual field and beyond the field of view of the display devices in use. Mobile AR systems allow users to interact with all of the environment, rather than being focused on a limited screen area. Hence, they allow interaction during visual search, tool acquisition and usage, or navigation. In emergency services or military settings, AR can cue users to dangers, obstacles, or situations in the environment requiring immediate attention. These many applications call for a general purpose interface technique to guide user attention while mobile to physical and virtual objects, labels, locations and other information populating a potentially cluttered physical environment.

Mobile AR interfaces present an interface challenge that can be characterized as follows: How can a mobile interface manage and guide visual attention to locations in the environment where critical information or objects are present, even when they are not within the visual field? The challenge is part of a larger need for *attention management* (Roel 2002) in high information bandwidth mobile interfaces.

To illustrate the benefits of management of visual attention in an AR system, consider the following application scenarios:

Telecollaborative spatial cueing. An emergency technician wears a camera and an AR HMD while collaborating with a remote doctor during a medical emergency. The

remote doctor needs to indicate a piece of equipment that the technician must use next. What is the quickest way to direct her attention to the correct tool among a large and cluttered set of alternatives, especially if she is not currently looking at the tool tray and doesn't know the technical term for the tool

Object Search. A warehouse worker uses a mobile AR system to manage inventory, and is searching for a specific box in an aisle where dozens of virtually identical boxes are stacked. Tracking systems integrated into the warehouse detect that the box is stored on a shelf behind the user using inventory records, a Radio Frequency Identification (RFID) tag, or other markers. What is the most efficient way to signal the target location to the user?

Procedural Cueing during Training. A trainee repair technician uses an AR system to learn a sequence of steps where parts and tools are used to repair complex manufacturing equipment. How can the computer best indicate which tool and part to grab next in the procedural sequence, especially when the parts and tools may be distributed 360° throughout a large workspace?

Spatial Navigation. A tourist with a Personal Digital Assistant (PDA) equipped with Global Positioning System (GPS) is looking for an historic building in a street with many similar buildings. The building is around the corner down the street. How can the PDA efficiently indicate a path to the main entrance?

These scenarios share a common demand for a technique that allows for precise target location cueing in near or far open spaces and at any angle relative to the user under conditions where speed and accuracy may be important. Any technique must be

able to provide continuous guidance and direct the user around occlusions. The scenarios illustrate various cases where attention must be guided or managed by the interface.

7.1 Attention Management

Human cognitive capacity is a finite resource and attention is one of the most limited of mental resources (Shiffrin 1979). Attention management (Roel 2002) is a key human-computer interaction issue in the design of interfaces and devices (Horvitz, Kadie, Paek and Hovel 2003; McCrickard and Chewar 2003). Information-rich applications of mobile AR interfaces (e.g., emergency services) begin to push up against a fundamental human factors limitation, the limited attention capacities of humans. For example, the attention demands of relatively simple and low bandwidth mobile interfaces, such as PDAs and cell phones, may contribute to car accidents (Redelmeier and Tibshirani 1997; Strayer and Johnston 2001).

Attention is used to focus cognitive capacity on a certain sensory input so that the brain can concentrate on processing the information of interest (van der Heijden 1992; van der Heijden 2003). Attention is primarily directed internally, from the “top down” according to the current goals, tasks, and larger dispositions of the user. Attention, especially visual attention, can also be cued by the environment. For example, attention can be user driven, i.e., “find the screwdriver,” collaborator driven “use this scalpel now,” or system driven “please use this tool for the next step.”

Visual attention is even more limited, since the system may have information about objects anywhere in an omnidirectional working environment around the user. Visual attention is limited to the field of view of human eyes ($<200^\circ$), and this limitation is further narrowed by the field of view of common HMDs ($<80^\circ$).

In mobile AR interfaces the attentional demands of the interface on mental workload (Hancock and Meshkati 1988; Johnson and Proctor 2004) must also be considered. Attention is shared across many tasks and tasks in the virtual environment are often not of primary consideration to the user. Individuals may be ambulatory, working with physical tools and objects, and interacting with others. The user may not be at the correct location in the scene, or looking at the correct spatial location or object needed to accomplish a task. So, attention management in the interface should reduce demands on mental workload.

7.1.1 Attention Cueing in Existing Interfaces

Currently, there are few, if any, general mobile interface paradigms to quickly direct spatial attention to objects or locations anywhere in the environment. Users and interface designers have evolved various ways to direct visual attention in interpersonal interaction, architectural settings, and standard interfaces.

7.1.1.1 Spatial cueing in Windows Interfaces

WIMP interfaces benefit from the assumption that user's *visual attention is directed to the screen*, which occupies a limited angular range in the visual field. Visual cues such as flashing cursors, pointers, radiating circles, jumping centered windows, color contrast, or content cues are used to direct visual attention to spatial locations on the screen surface. Large display areas extend this angular range, but still limit the visual attention to a clearly defined area. Khan and colleagues (Khan, Matejka, Fitzmaurice and Kurtenbach 2005) proposed a visual spotlight technique for large room interfaces.

The integration of audio with visual cues helps draw attention even when vision is not directed to the screen. Of course, these systems work within the confines of a very

limited amount of screen real estate; an area most users can scan very quickly. The audio cue often only initiates the attention process, requiring completion using visual scanning. These techniques cannot easily or quickly cue objects in the 3D environment around the user, for example pointing at an object behind the user.

7.1.2 Spatial Cueing in Augmented Reality

In mobile AR environments, the volume of information is large and omnidirectional. AR environments have the capacity to display large amount of informational cues to physical objects in the environment.

Most current AR systems adopt WIMP cursor techniques or visual highlighting to direct attention to an object (e.g., Feiner *et al.* 1993; Mann 2000). Recently, Chia-Hsun and colleagues (Bonanni, Lee and Selker 2005) proposed projecting light into the environment. Other techniques involve adding virtual quasi-architectural signage or virtual objects such as arrows or lines to the environment (Schmalstieg and Wagner 2005).

Spatial cueing techniques used in interpersonal communication (Burgoon, Buller and Woodall 1996), WIMP interfaces, and architectural environments are not easily transferred to AR systems. Almost all of these techniques assume that the user is looking in the direction of the cued object or that the user has the time or attentional capacity to search for a highlighted object. Multimodal cues such as audio can be used to cue the user to perform a search, but the cue provides limited spatial information and must compete with other sound sources in environment. Spatialized audio (Baluert 1983) does not have the spatial resolution to indicate spatial locations precisely.

7.2 The Omnidirectional Attention Funnel

Interface design in a mobile AR system presents two basic challenges in managing and augmenting attention of the user:

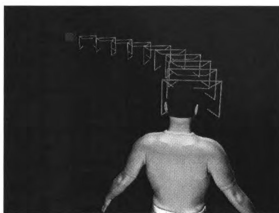


Figure 7.1. The attention funnel links the head of the viewer directly to an object anywhere around the body.

(1) Omnidirectional cueing. To quickly and successfully cue visual attention to any physical or virtual object in 360° space as needed.

(2) Minimal attention demands. Minimize mental workload and attention demands during search or interference with attention to tasks, objects, or navigation in the physical environment.

The Omnidirectional Attention Funnel is an AR display technique for rapidly guiding visual attention to any location in physical or virtual space. The basic components of the attention funnel are illustrated in Figure 7.1. The most visible component is the set of dynamic 3D virtual objects linking the view of the user directly to the virtual or physical object.

In spatial cognitive terms, the attention funnel visually links a head-centered coordinate space directly to an object centered coordinate space, funneling focal spatial attention of the user to the cued object. The attention funnel takes advantage of spatial cueing techniques impossible in the real world, and AR's ability to dynamically overlay 3D virtual information onto the physical environment.

Like many AR components the AR funnel paradigm consists of: (1) a display technique, the attention funnel, combined with (2) methods for tracking and detecting the location of objects to be cued.

7.2.1 Components of the Attention Funnel

The attention funnel has been realized as an interface widget in an augmented reality development environment. The attention funnel interface component (*arwattention*) and is one component in a planned set of user interface *widgets* being designed for mobile AR applications. These components are being built and tested as extensions of the ImageTclAR augmented reality development environment (Owen *et al.* 2003). The *arwattention* widget provides a mechanism for drawing visual attention to locations, objects, or paths in an AR environment.

The basic components of the attention funnel, as illustrated in Figure 7.2, are: (a) a view plane pattern with a virtual boresight in the center, (b) a dynamic set of attention funnel planes, (c) an object plane with a target graphic, and (d) a invisible curved path linking the head or viewpoint of the user to the object. Along this path are placed patterns that are repeated in space and normal to the line. We refer to the repeated patterns on the linking path as an attention funnel.

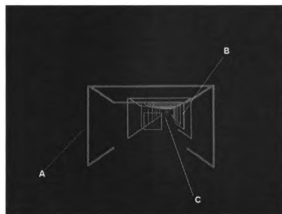


Figure 7.2. Three basic patterns are used to construct a funnel: (A) the head centered plane includes a boresight to mark the center of the pattern from the user's viewpoint, (B) funnel planes, added in a fixed pattern (approximately every 12 centimeters) between the user and the object, and (C) the object marker pattern that includes a red cross hairs marking the approximate center of the object.

The path is defined using cubic curve segments. Initial experiments have instantiated the path as Hermite curve (Hearn and Baker 1996). A Hermite curve is a cubic curve segment defined by a start location, end location, and tangent vectors at each end. The curve follows a path from the starting point in the direction of the starting end tangent vector. It ends at the end point with the curve approaching the end point in the direction of the end tangent vector. As a cubic curve segment, the curve presents a smoothly changing path from the start point to the end point with curvature controlled by the magnitude of the tangent vectors. Hermite curves are a standard cubic curve method discussed in any computer graphics textbook. Figure 7.3 clearly illustrates the curvature of the funnel from a bird's eye perspective.

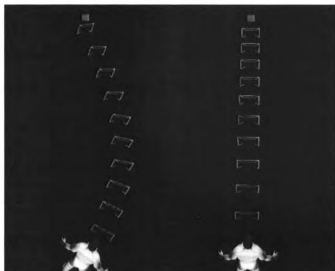


Figure 7.3. As the head and body move, the attention funnel dynamically provides continuous feedback. Affordances from the perspective cues automatically guide the user towards the cued location or object. Dynamic head movement cues are provided by the skew (e.g., left, right, up, down) of the attention funnel. The level of alignment (skew) of the funnel provides an immediate intuitive sense of how much the body or head must turn to see the object.

The starting point of the Hermite curve is located at some specified distance in front of the origin in a frame defined to be the viewpoint of the user (the center of projection for a single viewpoint or average of two viewpoints for stereo viewers). The curve terminates at the target. The tangent vector for the Hermite curve at the starting point is in the $-z$ direction¹ and the tangent vector at the ending point is a vector specified as the difference between the end and start locations (the direction to the target). The curvatures of the starting and ending points are specified in the application.

¹ Assuming a right hand coordinate system.

A single cubic curve segment creates a smoothly flowing path from the user's viewpoint to the target in a near field setting. Larger environment that include occlusions are require complex navigation are realized using a sequential set of cubic curve segments. The join points of the curve segments are specified by a navigation computation that takes into account paths and occlusions. As an example, a larger outdoor navigation system under development uses the Microsoft® Mappoint® commercial map management software to compute waypoints on a navigation path that then serve as the curve join points for the attention funnel path. The key design element is the smooth curvature of the path that allows for the funneling of attention in the desired target direction.

The orientation of each pattern along the visual path is obtained by spherical linear interpolation of the up direction (Shoemake 1985). Spherical interpolation allows the rotation angle between each interval to be constant, i.e. the changes of orientations of the patterns are smooth. The computational cost of this method is very small, involving the solution of the cubic curve equation (three cubic polynomials), the spherical interpolation solution, and computation of a rotation matrix for each pattern display location. Computational costs are dwarfed by the rendering costs for even this low-bandwidth display rendering.

The purpose of an attention funnel is to draw attention when it is not properly directed. When the user is looking in the desired direction, the attention funnel becomes superfluous and can result in visual clutter and distraction. The solution to this case is to fade the funnel as the dot product of the source and target tangent vectors approaches one, indicating the direction to the target is close to the view direction.

7.2.2 Affordances in the Attention Funnel that Guide Navigation and Body Rotation

The attention funnel uses various overlapping visual cues that guide body rotation, head rotation, and gaze direction of the user.

Building on an attention sink pattern introduced by Hochberg (Hochberg 1986), the attention funnel uses strong **perspective cues** as shown in Figure 7.4. Each attention funnel plane has diagonal vertical lines that provide depth cueing towards the center of the pattern. Each succeeding funnel plane is placed so that it fits within the preceding plane when the planes are aligned in a straight line. Increasing degrees of alignment cause the interlocking patterns to draw visual attention towards the center. Three basic patterns are used to construct a funnel: (1) the head centered plane includes a bore sight to mark the center of the pattern from the user's viewpoint, (2) funnel planes, added in a fixed pattern (currently every 12cm) between the user the object, and (3) the object marker pattern that includes a red bounding box marking the approximate center of the object. Patterns 1 and 3 are used for dynamically cueing the user that they approach an angle where they are "locked onto" the object (see below).

As the head and body moves, the attention funnel provides continuous feedback that indicates to the user how to turn the body and/or head towards the cued location or object. Continuous dynamic head movement cues are indicated by the skew (e.g., left or right) of the attention funnel. The pattern of the funnel provides an immediate intuitive sense of the location of object relative to the head. For example, if the funnel skews to the right, the user knows to move his head to the right (e.g., more skewing suggests that more body rotation is needed to see it). The funnel provides a continuous dynamic cue that one is getting closer to being "in sync" and locked onto the cued object. When looking

directly at the object, the funnel fades so as to minimize visual clutter. A target behind the user is indicated by a funnel that moves forward for visibility, then turns and heads behind the user - a clear visual cue.

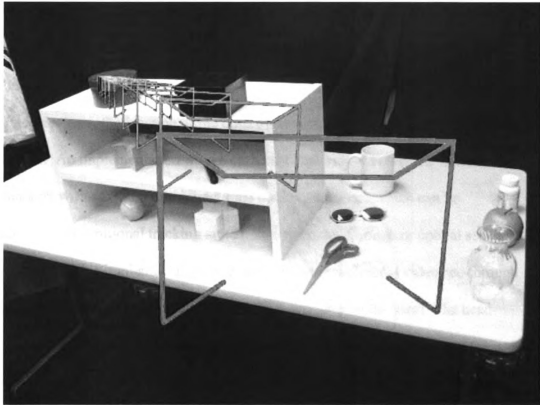


Figure 7.4. Example of the attentional funnel drawing attention of the user to an object on the shelf, the red box.

7.2.3 Methods for Sensing or Marking Targets Objects or Locations

Attention funnels are applicable to any augmented vision display technology capable of presenting 3D graphics, including head-mounted displays and video see-through devices such as tablet PC's or handheld computers. The location of target objects or locations in the environment may be known to the system because they are: (1) virtual

objects in tracked three-dimensional space, (2) tagged with sensors such as visible markers or RFID tags, or (3) at predefined spatial locations as in GPS coordinates. Virtual objects in tracked 3D space are the most straightforward case, as the attention funnel can link the user to the location of the target virtual object dynamically. Objects tagged with RFID tags are not necessarily detectable at a distance or locatable spatially with a high degree of accuracy, but local sensing in a facility may be sufficient to indicate a position sufficient for attention direction.

In some cases, the location of the object is detected by sensors and is not known ahead of time. An implementation we are currently exploring involves the detection of visible markers with auxiliary omnidirectional tracking cameras, which can be implemented as an additional tracking system in a video see-through or optical see-through system. (This implementation is distinct from the traditional video see-through system, where the only camera used represents the viewpoint of the user). The head-mounted omnidirectional camera detects markers in a 360° environment around the user. The relation of the camera to the user's viewpoint is known. Detected objects can be cued for the user based on task needs or search requests by the user (i.e., "find the tool box").

7.3 Methodology

A within-subjects experiment was conducted to test the performance of the attention funnel design against other conventional attention direction techniques: visual highlighting and verbal cues. The experiment had one independent variable, the method used for directing attention, with three alternatives: (1) the attention funnel, (2) visual highlight techniques, and (3) a control condition consisting of a simple linguistic cue.



Figure 7.5. Test Environment: The user sat in the middle of test environment for the visual search task. It consisted of an omnidirectional workspace assembled from four tables each with 12 objects (6 primitive shapes and 6 general office objects) for a total of 48 target search objects.

7.3.3 Apparatus and Test Environment

A 360° omnidirectional workspace was created using four tables as shown in Figure 7.5. 12 objects were placed on each table: 6 primitive objects of different colors (e.g. red box, or black sphere) on a shelf, and 6 general objects (e.g. stapler, notebook) on the table top.

Visual cues were displayed in stereo with the Sony Glasstron LDI-100B head-mounted display, and audio stimulus materials were presented with a pair of headphones. Head motion was tracked by an Intersense IS-900 ultrasonic/inertia hybrid tracking

system. Stereo graphics were rendered in real time based on the data from the tracker. A pressure sensor was attached to the thumb of a glove to capture the reaction time when the subject grasped the target object.

Presentation of stimulus materials, audio instructions for participants, experimental procedure sequencing, and data collection for the experiment was automated so that the experimenter did not need to manually record the experimental results. The experiment was developed in the ImageTclAR AR development environment (Owen *et al.* 2003).

7.3.4 Measurements

Search Time, Error, and Variability. Search time in milliseconds was measured as the time it took for participants to grab a target object from among the 48 objects following the onset of an audio cue tone. The end of the search time was triggered by the pressure sensor on the thumb of the glove when the user touched the target object. An error was logged for cases when participants selected the wrong object.

Mental Workload. Participant's perceived task workload in each condition was measured using the National Aeronautics and Space Administration Task Load Index (NASA TLX) after each experimental condition (Hart and Staveland 1988).

7.3.5 Procedure

Participants entered a training environment where they were introduced and trained to use each interface (audio, visual highlight, attention funnel). They then began the experiment. Each subject experienced the interface-treatment conditions (audio, visual highlight, and attention funnel) in a randomized order. For each condition, participants were cued to find and touch one of the 48 objects in the environment as

quickly and accurately as possible. Participants participated in 24 trials balanced such that 12 trials involved searching for a random selection of primitive objects and 12 trials involved randomly selected general everyday objects.

7.4 Results

A general linear model repeated measure analysis was conducted to test the effect of metaphors on the different performance indicators. There was a *significant effect* of interface type on search time, $F(2, 14) = 10.031$, $p = 0.001$, and on search time consistency (i.e., smallest standard deviation), $F(2, 14) = 23.066$, $p = 0.000$. The attention funnel interface clearly allows subjects to find objects in the least amount of time and with the most consistency ($\underline{M} = 4473.75$ ms, $\underline{SD} = 1064.48$) compared to the visual highlight interface ($\underline{M} = 6553.12$, $\underline{SD} = 2421.10$) and the audio only interface ($\underline{M} = 4991.94$ ms, $\underline{SD} = 3882.11$), which had the largest standard deviation. See Figure 7.6.

There was a significant effect of interface type on the participants perceived mental workload, $F(2, 14) = 4.178$, $p = 0.027$. The results indicate that the attention funnel interface has the lowest mental workload ($M = 44.64$, $SD = 16.96$), comparing to the visual highlight interface ($M = 54.57$, $SD = 18.26$), and the audio interface ($M = 55.57$, $SD = 12.43$). See Figure 7.7.

There was no significant effect of interface type on error, $F(2, 14) = 1.507$, $p = 0.24$ (attention funnel $\underline{M} = 1.14$, $\underline{SD} = 0.77$, visual highlight, $\underline{M} = 1.43$, $\underline{SD} = 1.56$, audio $\underline{M} = 0.86$, $\underline{SD} = 1.03$).

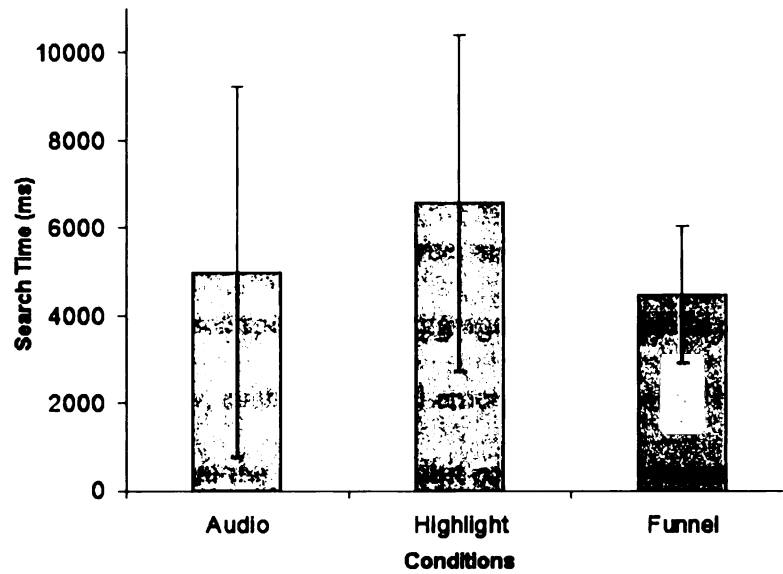


Figure 7.6 Search time and consistency by experimental condition. Attentional funnel decreased search time by 22% on average (28% when reach time is subtracted) and increased search consistency (decreased variability) by 65%.

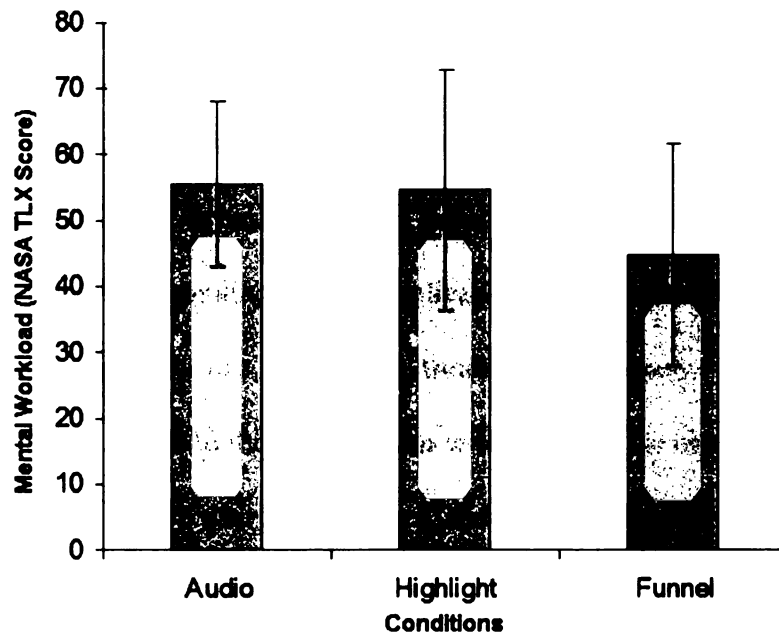


Figure 7.7. Mental workload measured by NASA TLX for each experimental condition.

7.5 Discussion

When compared to standard cueing techniques such as visual highlighting and audio cueing, we found that the attention funnel **decreased the visual search by 22%** overall, or approximately 28% for visual search time, and 14% over its next fastest, as shown in Figure 6. While increased speed in the aggregate is valuable in some applications of augmented reality, such as medical emergency and other high risk applications, it may be critical that the system exhibit consistent performance. The attention funnel had a very robust effect on search consistency (decreased standard error). The interface **increased consistency by 65%** on average, and 56% over the next best interface.

In summary the attention funnel led to faster search and retrieval times, greater consistency of performance, and decreased mental workload when compared to verbal cueing and visual highlighting techniques.

7.6 Application of the Attention Funnel

With the success of AR enabled, mobile systems, designers will seek to add potentially rich, even unlimited layers of location based information onto physical space. As AR systems are used in demanding mobile applications such as manufacturing assembly, warehouse search, tourism, navigation, training, and distant collaboration, interface techniques appropriate to the AR medium will be needed to manage the mobile user's limited attention, improve user performance, and limiting cognitive demands while achieving a more optimal spatial performance.

The attention funnel paradigm involves basic techniques that have potentially general applicability in mobile interfaces: A user's attention has to be directed to objects

or locations in order to accomplish tasks. We are currently implementing the technique on other mobile devices including hand held devices such as PDAs and cell phones.

Broadly, the attention funnel techniques may be implemented in applications involving the following generic classes of fundamental tasks:

Physical Object Selection. Situations where a user may be looking for a physical object in a space, for example a tool in a workbench, a box in a warehouse, a door in space, the next part to assemble during object assembly, etc. The system can direct the user to the target object.

Virtual Object Selection. AR systems may insert labels or 3D objects inside the environment. These may be within or outside the current view of the user. Attention funnels can cue them to look at the spatially registered label, tool, or cue.

Visual Search in a Cluttered Space. The user may be searching in a highly cluttered natural or artificial environment. An attention funnel can be used to cue them to the correct location to view, even if they are not looking in the right place.

Navigation in Near Space. The system might also need to direct the walking path of the individual through near space (e.g., through aisles, etc.). A directional funnel path (slightly different implementation than the attention funnel above) can be used to indicate and cue the user's direction, and provide dynamic cues as to path accuracy.

Navigation in Far Space. An attention funnel can direct users to distant landmarks. As an example, someone walking towards an office several blocks away must maintain a link to the landmark as they navigate through an urban environment, even when landmarks are obscured.

The AR attention funnel paradigm represents an example of cognitive augmentation specifically adapted for users of mobile AR systems navigating and working in information and object rich environments.

8 Discussion and Conclusion

This thesis is the first research work to construct a spatial framework of information placement in AR environments based on neuropsychological research. The spatial framework provides a theoretical model to map cognitive properties of each Infospace for information organization in mobile AR interface. Unique cognitive properties of each infospaces in the spatial framework is then systematically review. A large volume of literature in psychology, behavioral science and neuroscience on spatial cognition is systematically reviewed and organized according the spatial framework. As there is no spatial framework previously existed for information placement in three-dimensional space, compiling a set of cognitive properties in each infospace in the framework allows researchers to determine where new research is needed to further investigate issues in AR interfaces design.

Three research questions were identified concerning unexplored cognitive properties in the spatial framework that are useful for the design of AR environments. The first research question addresses the capability of the human cognitive system to manipulate egocentric and allocentric reference frames to encode spatial information in the environment. Even though there is no physical equivalent of Peripersonal Infospace in the real world, experimental results show that participants' preference of reference frame for spatial memory and spatial updating can be easily manipulated by oral instruction or brief experience. The ease of manipulation in spatial reference frame shows promise for using the Peripersonal space for information organization.

The second research question is regarding the applicability of perceptual asymmetric properties to information display in Egocentric Infospaces. There are a large number of perceptual asymmetric properties in the psychology literature concerning different cognitive properties such as reaction time, emotion and semantic meaning. Two experiments were conducted to evaluate the applicability of these asymmetric properties to AR interfaces design. Results show that perceptual asymmetric properties based on reaction time (e.g. perceptual response, memory retrieval), typically measured in milliseconds, are too subtle to have practical impact on reaction to stimuli in AR and other information displays, and should only be used sparsely for information placement in egocentric infospaces. On the other hand, semantic meaning and participants' emotional response of virtual objects and agents change with the location around the body. AR interface designers may need to be aware that location may impart connotation and emotion to user's perception of the information objects, especially agents.

Finally, a novel metaphor for directing visuo-spatial attention, the Attention Funnel, was developed. Traditional paradigms for directing user's attention (such as blinking, audio signal, audio instruction, use of color and highlight) are inaccurate, mentally demanding and ambiguous in the omni-directional environment. The Attention Funnel paradigm, a dynamic three-dimensional perspective cue linking user's retinotopic space to a virtual or physical object in space, was shown to reduce user's visual search time and mental workload comparing with traditional paradigms.

8.1 Guideline for Information Display in Augmented Reality Environments

The cognitive properties in the spatial framework developed form the basis of information display guideline in AR environments. Based on the discussion in this

thesis, a set of guidelines for information display in mobile AR interfaces was compiled (Appendix A). This set of guidelines are the distillation of the best available information about behavioural patterns in spatial cognitive psychology and results of new experimentation presented in this thesis, rather than general rules of thumb derived from user or designer experiences and/or personal opinions. The set of guidelines provide interface designers a clear framework to follow for spatial information placement in AR environments, helping them to develop AR interfaces that exploit the spatial processing capabilities of the human brain. It also serves as a checklist for AR interface designers, guiding decisions on a wide range of issues in spatial information placement during the design and evaluation process. More importantly, this thesis seeks to promote discussion among researchers and stimulate additional research in spatial AR interfaces design.

8.2 Future Works

As there is no clearly defined guidelines previously existed for spatial information placement in three-dimensional space, compiling a set of initial guidelines allows researchers to determine where new research is needed to further investigate issues in AR interfaces. The compiled set of guidelines in Appendix A is by no mean a complete set of guidelines that covers every aspect of spatial information placement in AR environments. Issues raised in this dissertation only serve as a starting point, and it is expected that AR Interfaces research to use this set of guideline to identify new issues not being covered in this guideline and expanding this set of guidelines with new experimental results.

There is no physical equivalent of Peripersonal Infospace in the real world, and precious little about cognitive properties of Peripersonal Infospace is known. For example, spatial location in Peripersonal Infospace may interact with many other cognitive processes such as information search, attention, visual change detection and memory. More research in cognitive psychology related to the Peripersonal Infospace is needed to explore cognitive properties of the Peripersonal Infospace and turn these properties into spatial information display guideline.

Another under explored Infospace in the spatial framework is the Personal-body Infospaces. There were just a few studies in neuroscience exploring the relation of tool usage and personal space and body schema in the last few years. But these studies are limited to basic research in neuroscience. There are a lot of unexplored cognitive properties related to Personal-body Infospace (such as reaching response, memory, accuracy, emotion, semantic meaning) that have potential implications in AR interface design.

A prototype mobile AR interface based of the compiled set of guidelines is currently being developed. It is intended to be developed for mobile AR applications field experiment in various task specific applications and scenario. It will also be used as a test bed for future studies to explore new issues in mobile AR interfaces.

8.3 Conclusion

The design of AR interfaces prompts a significant human factors challenge of mapping different metaphors, information, and functions of computer usage into the human cognitive system. In this thesis, a set of spatial information placement guidelines was constructed based on a body of literature in neuropsychology, spatial cognition, and

behavioral sciences and a series of experiments tightly related to AR interfaces design. The literature and experimental results provides grounding for theory driven human-computer interaction design for the development of high performance AR interfaces, mobile infospaces potentially tailored to human spatial cognition.

The set of guidelines present interface designers a clear framework to follow for spatial information placement in AR environments, and serves as a checklist for AR interface designers, guiding decisions on a wide range of issues in spatial information placement during the design and evaluation process. More importantly, this thesis seeks to provoke discussions among researchers and stimulate more research in spatial AR interfaces design.

Appendix A. Spatial Information Display Guideline for Mobile Augmented Reality Interfaces

The guidelines presented in this chapter are the distillation of the best available information about behavioural patterns in spatial cognitive psychology along with results of new experimentation. The guidelines present interface designers a clear framework to follow for spatial information placement in AR environments, helping them to develop AR interfaces that exploit the spatial processing capabilities of the human brain. It also serves as a checklist for AR interface designers, guiding decisions on a wide range of issues in spatial information placement during the design and evaluation process. More importantly, this thesis seeks to promote discussion among researchers and stimulate additional research in spatial AR interfaces design. As no clearly defined guidelines previously existed for spatial information placement in three-dimensional space, compiling a set of initial guidelines for spatial information placement allows researchers to determine where new research is needed to further investigate issues in AR interfaces.

The compiled set of guidelines in this thesis only addresses spatial issues in AR interfaces design. It presents a set of issues related to how the human brain processes information objects in space in a mobile AR computing environment. This set of guidelines does not address general interface and display issues such as styles, appearances, messages and contents. It is also expected that additional spatial interface guidelines will be advanced as research into human spatial cognition proceeds, thereby expanding this set of guidelines with new ideas. This is, by no means, the end of research into spatial placement guidelines for augmented reality systems. Rather, it is intended as a strong beginning.

A. Spatial Framework of Three-dimensional Space

A1. Partitioning Three-dimensional Space into Information Spaces

Guideline:	AR systems should include in their design the inherent partitioning of space into five <i>infospaces</i> , supporting each as an identifiable information frame.
Comments:	<p>The five infospaces are defined as: (a) Personal-body Infospace, (b) Peripersonal Infospace, (c) Extrapersonal Focal Infospace, (d) Extrapersonal Action-Scene Infospace, and (e) Extrapersonal Ambient Infospace. This partitioning of space is supported by neuropsychological research that shows these spaces to have unique physical and psychological properties. Information objects should be categorized as members of an appropriate infospace.</p> <p>The definition of each Infospace is given as follow: (a) Personal-body Infospace is the volume immediately adjacent to and including the surfaces of the user's body, (b) Peripersonal Infospace is the volume defined by the arm-reaching space immediately in front of the body, (c) Extrapersonal Focal Infospace is an elliptical region with a lateral extent of 20°-30° anchored to the user's eye fixation. This is the predominant visual space and is the target space for head stabilized reference frames, although the best definition of extrapersonal focal spaces would be based on eye tracking information, (d) Extrapersonal Action-scene Infospace is the spatial volume of the allocentrically oriented</p>

	spaces. It encapsulates the body in a 360° surround, with a range starting from 2 meters from the body to approximately 30 meters, and (e) Extrapersonal Ambient Infospace is the earth-fixed outermost space of the visual field.
--	--

A2. Egocentric Infospaces

Guideline:	Interface elements that require access regardless of the user's location should be attached to any of the three egocentric Infospaces: Personal-body Infospace, Peripersonal Infospace, and Extrapersonal Focal Infospace; unless the volume required exceeds the capacity of the Peripersonal Infospace
Comments:	Interface elements attached to the three egocentric Infospaces are reachable by a mobile user during locomotion regardless of location. However, egocentric infospaces have limited capacity. Information objects that require a volume exceeding the capacity of the Peripersonal Infospace, which has the largest capacity among the three egocentric Infospaces, should be attached to the Extrapersonal Action-scene Infospace.

B. Peripersonal Infospace

B1. Reference Frames and tracking requirements of Peripersonal Infospace

Guideline:	Tracking for the peripersonal infospace should be attached as
-------------------	---

	closely as possible to the spine, ideally to the upper lumbar vertebrae.
Comments:	Information objects in the <i>Peripersonal Infospace</i> remain stationary with respect to the upper torso. Tracking of the upper back (dorsal area) creates a frame for information objects attached to the peripersonal infospace that follows the body, but without the unwanted breath motion exhibited by the breast.

B2. Physical Volume and Information Capacity of Peripersonal Infospace

Guideline:	The Peripersonal Infospace is the default Infospace for information objects that must follow the user during locomotion and interface elements that require frequent access.
Comments:	The Peripersonal Infospace has the highest capacity, and exhibits the least psychophysiological specialization among the three egocentric Infospaces. It is ideal as the default Infospace for generic interface elements that must follow the user during locomotion and interface elements that require frequent access. A general rule is that information objects that must follow the user during locomotion should be initially assigned to the Peripersonal Infospace unless some unique characteristic of the other two egocentric spaces demands their application.

B3. Visibility of Peripersonal Infospace

Guideline:	Interface designs must accommodate the variable visibility of differential spatial locations in Peripersonal Infospace.
Comments:	<p>The space immediately in front of the head is the most visible volume in Peripersonal space. Information object visibility decreases as they are moved farther away from the central area.</p> <p>Designers should map the visibility of spatial location as elements of the design, ensuring that objects that require more attentiveness are placed at the location with higher visibility.</p>

B4. Spatial Bias of Peripersonal Infospace

Guideline:	Placement of interface elements should be spatially sorted so as to accommodate the spatial biases of the Peripersonal Infospace.
Comments:	<p>Reaction time of reaching movements is biased towards the lower portion of the Peripersonal Infospace and the middle 60° of the body. Hand motor resolution is also finer in the lower portion of Peripersonal Infospace. There are evidence that these motor advantages extend into a memory advantage for recalling the location of objects and recognition of objects that are manipulated.</p> <p>For right-handed users, these properties are also biased towards the right side.</p>

C. Personal-body Infospace

C1. Reference Frames and tracking requirements of Personal-body Infospace

Guideline:	AR user interfaces may incorporate multiple Personal-body Infospaces by tracking alternate body parts.
Comments:	A Personal-body Infospace is a reference frame stabilized to a body part such as the hand or arm. Information objects in a Personal-body Infospace remain stationary with respect to the body part they are attached to. Different body parts lead to different capabilities. Humans are used to associating information with the arm, but less is know about the association of information objects with other body parts.

C2. Physical Volume and Information Capacity of Personal-body Infospace

Guideline:	The amount of information placed in a Personal-body Infospace should be limited in any design.
Comments:	The Personal-body Infospace has a very limited volume with the smallest capacity of any of the three Egocentric Infospaces. The capacity of a Personal-body Infospace depends upon the area of the space surrounding the body part. The volume of a Personal-body Infospace typically extends a few centimeters from the epidermis. However, there is neuropsychological evidence that the Personal-body Infospace can be plastically extended following active tool-use. So the volume can possibly be extended to the surrounding volume of interface elements attached to Personal-body Infospace after prolonged active use of those interface

	elements.
--	-----------

C3. Visibility of Personal-body Infospace

Guideline:	Interface designs must accommodate the variable visibility of objects in a personal-body infospace.
Comments:	Information objects in Personal-body Infospace are not always visible to the user. Information objects attached to the forearms are the most visible, while information objects attached to the upper torso, lower torso, upper-arms, thighs and legs are less visible due to the limits of head motion. Designers should map the visibility of these objects as elements of the design, ensuring that objects that must remain visible are not placed in regions commonly occluded or beyond the normal field of view. Some user interface elements do not require visibility, such as physical interfaces like buttons, but are ideally associated with a Personal-body Infospace because of the physical reachability of objects attached to the infospace.

C4. Spatial Bias of Personal-body Infospace

Guideline:	Interface design should take spatial bias into consideration during the selection of the appropriate Personal-body Infospace for a given interface element.
Comments:	Personal-body Infospaces are strongly biased towards the ventral body (frontal region). The dorsal body (back of the body) is not

	<p>within the visual field and less accessible by the user's hands. The Personal-body Infospace is further biased towards the upper body, where body parts are reachable by the hands.</p>
--	--

C5. Proprioception

Guideline:	<p>Tasks and control functions that require a high degree of spatial resolution should be attached to one of the Hand-stabilized Personal-body Infospaces in order to take advantage of proprioceptive feedback for high accuracy placement.</p>
Comments:	<p>Proprioception, the sensation of the movement and orientation of body parts, is required to achieve high accuracy hand manipulation and alignment. A Personal-body Infospace associated with the hands provides the best proprioceptive feedback to the user.</p>

D. Extrapersonal Focal Infospace

D1. Reference Frames and Tracking Requirement of Extrapersonal Focal Infospace

Guideline:	<p>Tracking of head and/or eye motion is required for use of the extrapersonal focal infospace.</p>
Comments:	<p>Information objects in the Extrapersonal Focal Infospace remain stationary with respect to eye fixation, or with respect to the user's head when eye tracking is not available. The head is traditionally tracked in AR systems, typically through tracking of a head-mounted display that is assumed to remain fixed relative to the</p>

	head. An eye movement tracker is required for information to remain stationary with respect to eye fixation.
--	--

D2. Physical Volume and Information Capacity of Extrapersonal Focal Infospace

Guideline:	The amount of information placed in the Extrapersonal Focal Infospace should be limited in any design.
Comments:	The physical volume of the Extrapersonal Focal Infospace is the volume immediately in front of the head, and its capacity for information objects is necessarily very limited. Furthermore, the central area should be reserved to avoid visual clutter that may obscure the real environment.

D3. Visibility of Extrapersonal Focal Infospace

Guideline:	Information objects that require immediate attention should be attached to Extrapersonal Focal Infospace.
Comments:	Information objects in Extrapersonal Focal Infospace are located immediate in front of the head, and are always visible by definition regardless of the user's location and posture. Consequently, information objects in the Extrapersonal Focal Infospace have a great potential for distraction or interference with vision.

D4. Perceptual Fading of Visual Stimulus in Head Stabilized Reference Frame

Guideline:	Information objects that require sustained attention should not be placed in the Extrapersonal Focal Infospace.
Comments:	When attaching information objects to the Extrapersonal Focal Infospace, interface designers should be aware of perceptual fading, which may cause information to perceptually disappear over a period of time, ranging from seconds to minutes.

D5. Visual Clutter and Spatial Bias of Extrapersonal Focal Infospace

Guideline:	Even though user attention is biased towards the central area of the Extrapersonal Focal Infospace, information objects should be placed along the peripheral area of the Extrapersonal Focal Infospace to avoid visual clutter that may obscure the real environment.
Comments:	Visual attention is eccentrically distributed from the eye fixation point. However, information objects placed within a 5 degree radius of the eye fixation will cause annoyance to the user. Therefore, information objects should be placed at the peripheral area of the head-stabilized Extrapersonal Focal Infospace.

D6. Directing User's Attention in an Omni-directional Environment

Guideline:	The Attention Funnel paradigm, a dynamic three-dimensional perspective cue linking a user's retinotopic space to a virtual or physical object in space, is recommended for directing visuo-
-------------------	---

	spatial attention.
Comments:	Traditional paradigms for directing attention (such as blinking indicators, audio signals, audio instruction, use of color and highlighting) are inaccurate, mentally demanding and ambiguous in an omni-directional environment. The Attention Funnel paradigm, a dynamic three-dimensional perspective cue linking user's retinotopic space to a virtual or physical object in space, has been shown to reduce visual search time and mental workload comparing with traditional paradigms.

E. Egocentric Infospaces

E1. Spatial Asymmetric Properties in the Brain

Guideline:	Perceptual and kinematics asymmetric properties can be used to optimize the placement of information objects and interface elements in egocentric infospaces.
Comments:	Different quadrants in the visual field take a different visual pathway to different regions of the brain, and have different perceptual properties. Human motor skills are asymmetric due to cerebral lateralization. Presenting information on the correct side of the body could enhance the perceptual and cognitive processes relevant to the information objects, and result in a faster, more natural, and increasingly accurate access.

E2. Kinematics Asymmetric Properties: Unimanual Tasks

Guideline:	Simple pointing and selection tasks should be presented to the side of the dominant hand.
Comments:	Unimanual tasks are usually biased towards the dominant hand. The dominant hand is better at precise, corrective, and rapid movements.

E3. Kinematics Asymmetry Properties: Spatial Reference in Bimanual Tasks

Guideline:	Bimanual tasks that involve information objects requiring a physical stabilizing action, defined steady states, or a defining spatial reference should be placed on the non-dominant side.
Comments:	Motion of the dominant hand typically finds its spatial reference as the results of motion of the non-dominant hand. The roles of the non-dominant hand include physical stabilizing actions (e.g. stabilizing the container for precise selection and manipulation), defining steady states (e.g. aiming a moving target), and defining a spatial reference.

E4. Kinematics Asymmetry Properties: Spatial-temporal Scale for of the Two Hands

Guideline:	Information objects for bimanual interactions that require macrometric movement should be presented to the side of the non-dominant hand, and tasks that require micrometric movement should be presented to the side of the dominant hand.
Comments:	The dominant hand has a finer spatial and temporal motor

	resolution than the non-dominant hand.
--	--

E5. Kinematics Asymmetry Properties: Precedence in Action for the Two Hands

Guideline:	A single interface element that requires bimanual interactions should be presented on the side of the non-dominant hand.
Comments:	The non-dominant hand starts earlier than the dominant hand in bimanual action. Placing interface elements on the non-dominant side encourages reach and acquisition by the non-dominant hand.

E6. Perceptual Asymmetric Properties: Response time

Guideline:	Perceptual asymmetric properties based on reaction time (e.g. perceptual response, memory retrieval) are too subtle to have practical effects on reaction to stimuli in AR and other information displays, and should only be used sparsely for information placement in egocentric infospaces.
Comments:	There are a large number of perceptual asymmetric properties in the psychology literature. However, perceptual asymmetric properties based on reaction time are typically measured in milliseconds, and are too subtle to have significant effects on practical tasks.

E7. Perceptual Asymmetric Properties: Emotion

Guideline:	Information objects that intentionally trigger a user's emotion
-------------------	---

	should ideally be placed on the left side in Egocentric Infospaces.
Comments:	Information objects presented on the left side have semantic properties that are shown to deviate from neutral more significantly than objects falling on the right side. The left side of the body is more sensitive to stimulus-evoking emotions.

E8. Perceptual Asymmetries: Social Proxemics and Semantic Meaning

Guideline:	Information objects with conative meaning should ideally be placed closer to the body in an Egocentric Infospace.
Comments:	Information objects in the near space are perceived with more conative meanings (e.g. more relevant, superior, urgent, and aggressive), while information objects in the far space are perceived with less. This effect is stronger for agent representation such as representations of humans and animals.

F. Extrapersonal Action-scene Infospace

F1. Reference Frames of Extrapersonal Action-scene Infospace

Guideline:	Information objects can be attached to stationary objects in the environment without additional tracking support. Multiple Extrapersonal Action-scene Infospaces may be incorporated by tracking the motion of each moving object in the environment.
Comments:	Information objects in the Extrapersonal Action-Scene Infospace remain stationary relative to objects in the scene. Tracking

	<p>sources of typical AR systems induce a local reference frame.</p> <p>Information objects that remain stationary relative to stable objects in the environment can be pre-calibrated with respect to this local reference frame. Addition tracking is required for each moving object so that information objects remain stationary relative to the moving object.</p>
--	--

F2. Physical Volume of Extrapersonal Action-scene Infospace

Guideline:	Extrapersonal Action-scene Infospace can accommodate information objects that require a large volume.
Comments:	The physical volume of the Extrapersonal Action-scene Infospace is unlimited.

F3. Visibility of the Extrapersonal Action-scene Infospace

Guideline:	<p>Information objects attached to the Extrapersonal Action-scene Infospace require visuo-spatial directed attention paradigms (such as the Attention Funnel paradigm) should user's attention to the information object be necessary.</p>
Comments:	<p>Visibility of information objects in Extrapersonal Action-scene Infospace depends on the user's viewpoint orientation and position. Often this viewpoint and orientation will be such that the information object is beyond the field of view. When attention to the information object is required, visuo-spatial attention needs to</p>

	be directed explicitly.
--	-------------------------

F4. Remote Objects Selection in Extrapersonal Action-scene Infospace

Guideline:	Body parts used for pointing and selection tasks should be chosen based on the order of finger, hand, arm, and lastly, the head.
Comments:	Information objects may fall outside the reachable distance of the hands as the user navigates in the environment. In such cases, the object must be indicated by a direction indication rather than direct selection. This often entails pointing in the form of indicating a direction to the information object using the head, arm, hand, or finger. Performance of pointing tasks using the head has the highest Fitts's Index of Difficulty (i.e. most difficult) followed by the arm, the hand, and then the finger.

F5. Remote Objects Manipulation in Extrapersonal Action-scene Infospace

Guideline:	AR interface designers will need to analysis the requirements for of specific applications before choosing a remote manipulation technique. The pros and cons for various remote manipulation techniques are listed in Chapter 4 in Table 4.10.
Comments:	There is no standard remote objects manipulation technique that will work for all applications. Table 4.10 summarized the pros and cons of six remote manipulation techniques: (1) Raycasting, (2) CHIMP, (3) Arm-extension, (4) World in miniature, (5)

	HOMER, and (6) Voodoo Doll.
--	-----------------------------

G. Extrapersonal Ambient Infospace

G1. Reference Frames of Extrapersonal Focal Infospace

Guideline:	Information objects can be attached to stationary objects in the environment without additional tracking support.
Comments:	Information objects in the Extrapersonal Ambient Infospace remain stationary relative to objects on the earth. Once sufficient tracking is available to support Egocentric Infospaces, support for Extrapersonal Ambient Infospace is, for all practical purposes, free.

G2. Spatial Bias in Extrapersonal Ambient Infospace

Guideline:	Information objects should be placed nearer the floor in the Extrapersonal Ambient Infospace.
Comments:	The Extrapersonal Ambient Infospace is biased towards the peripheral visual field and lower visual field. Proximity to the floor provides a visual stabilization of the virtual elements relative to the real environment.

G3. Linear Perspective and Motion Perception Properties

Guideline:	The Extrapersonal Ambient Infospace is ideal for information objects related to spatial orientation and motion perception, for
-------------------	--

	examples, landmarks, horizontality cues and signage on the floor.
Comments:	Extrapersonal Ambient Infospace is particularly susceptible to linear perspective and optical flow cues.

H. Infospace Choice for Common Information Objects

H1. Alerts and System Messages

Guideline:	Alerts, system messages, or information objects that require immediate attention should be placed in the Extrapersonal Focal Infospace.
Comments:	The Extrapersonal Focal Infospace has the highest visibility among all Infospaces. For information objects attached to other Infospaces, user's immediate attention can be captured by an alert in the Extrapersonal Focal Infospace first. Devices such as the Attention Funnel can then be used to direct visuo-spatial attention to the location of the information objects.

H2. Unimanual Selection and Manipulation Tools

Guideline:	Selection tools and unimanual manipulation tools should be attached to the Personal-body Infospace of the dominant hand or fingers.
Comments:	Performance of pointing tasks using the hand or the finger has the lowest Fitts's Index of Difficulty. Furthermore, the dominant hand has a finer spatial and temporal motor resolution than the non-dominant hand.

H3. Tools Selection Tray

Guideline:	A <i>tools selection tray</i> should be attached to the Peripersonal Infospace.
Comments:	Peripersonal Infospace has the largest volume among the three egocentric Infospaces to accommodate various selection and manipulation tools. It also allows bimanual manipulation tools to be selected by both hands concurrently.

H4. Task Specific Information Objects related to the Real Environment

Guideline:	Task specific information objects related to the real environment should be place in the Extrapersonal Action-scene Infospace, and should be spatially registered to the task objects.
Comments:	The cost for information search and attention switching can be reduced by spatially placing task-related information in the correct spatial location.

H5. Information Objects that require continuous monitor

Guideline:	Information objects that require continuous monitoring (e.g. system or task specific statuses or readings) should be attached to the Peripersonal Infospace.
Comments:	Even though the Extrapersonal Focal Infospace is the most visible of the Infospaces, it is inappropriate for tasks that require sustained

	attention due to perceptual fading. The visibility of objects in a Peripersonal Infospace is acceptable for information that requires continuous monitoring.
--	--

H6. Non-task Specific System and Personal Information Storage: Small Volume

Guideline:	Peripersonal Infospaces and Personal-body Infospaces are ideal for non-task specific personal information objects that require a small volume.
Comments:	The metaphorical associations and proprioceptive memory established in the egocentric Infospaces provide for faster and more accurate access and manipulation of information objects. However, the number of information objects in an Extrapersonal Focal Infospace should be limited, as this space is not suitable for large volume system and personal information storage.

H7. Non-task Specific System and Personal Information Storage: Large Volume

Guideline:	Unregistered Extrapersonal Action-scene Infospace is ideal for holding non-task specific personal information objects that require a large volume.
Comments:	Extrapersonal Action-scene Infospace can accommodate information objects that require a large volume. System and personal information objects that exceed the capacity of Peripersonal Infospace can be attached to the unregistered Extrapersonal Action-scene Infospace.

9 References

- Abrahams, H., Krakauer, D. and Dallenbach, K. (1937). "Gustatory adaptation to salt." *American Journal of Psychology* 49(3): 462 - 469.
- Adamovich, S., Berkinblit, M., Hening, W., Sage, J. and Poizner, H. (2001). "The interaction of visual and proprioceptive inputs in point to actual and remembered targets in Parkinson's disease." *Neuroscience* 104(4): 1027 - 1041.
- Atkinson, J. and Egeth, H. (1973). "Right hemisphere superiority in visual orientation matching." *Canadian Journal of Psychology* 27: 152 - 158.
- Axelrod, S., Haryadi, T. and Leiber, L. (1977). "Oral report of words and word approximations presented to the left or right visual field." *Brain and Language* 1977: 550 - 557.
- Baluert, J. (1983). *Spatial hearing: the psychoacoustics of human sound localization*. Cambridge, MA, MIT Press.
- Bateson, G. (1972). *Steps to an ecology of mind*. New York, NY, Ballantine Books.
- Bauma, H. (1973). "Visual interference in the parafoveal recognition of initial and final letters of words." *Vision Research* 13: 767 - 782.
- Becklen, R. and Cervone, D. (1983). "Selective looking and the notice of unexpected events." *Memory & Cognition* 11: 601 - 608.
- Berti, A., Smania, N. and Allport, A. (2001). "Coding of far and near space in neglect patients." *Neuroimage* 14: S98 - S102.
- Biocca, F. (1997). "The cyborg's dilemma: Progressive embodiment in virtual environments." *Journal of Computer-Mediated Communication* 3(2).
- Biocca, F., David, P., Tang, A. and Lim, L. (2004). Does virtual space come precoded with meaning? Location around the body in virtual space affects the meaning of objects and agents. In *Proceedings of 54th Annual Conference of the International Communication Association*. New Orleans, LA. May 27 - 31, 2004.
- Biocca, F., Eastin, M. and Daugherty, T. (2001). Manipulating objects in the virtual space around the body: relationship between spatial location, ease of manipulation, spatial recall, and spatial ability. In *Proceedings of 51th Annual Conference of the International Communication Association*. Washington, DC. May 24 - 28, 2001.
- Biocca, F., Lamas, D., Gai, P., Brady, R. and Tang, A. (2001). Mapping the semantic asymmetris of virtual and augmented reality space. In *Proceedings of Fourth*

- International Conference on Cognitive Technology, CT2001*, 117 - 122. Warwick, UK. August 6 - 9, 2001.
- Biocca, F. and Rolland, J. (1998). "Virtual eyes can rearrange your body adaptation to virtual eye location in see-thru head-mounted displays." *Presence: Teleoperators and Virtual Environments* 7(3): 262-277.
- Bonanni, L., Lee, C. and Selker, T. (2005). Attention-based design of augmented reality interfaces. In *Proceedings of ACM CHI 2005*. Portland, OR. April 2 - 7, 2005.
- Boroditsky, L. and Ramscar, M. (2002). "The roles of body and mind in abstract thought." *Psychological Science* 13(3): 185.
- Bowman, D. and Hodges, L. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of ACM Symposium on Interaction 3D Graphics*, 35 - 38.
- Bratman, M. (1999). *Faces of intention: selected essays on intention and agency*. Cambridge, UK, Cambridge University Press.
- Brooks, F. (1996). "The computer scientist as toolsmith II." *Communications of the ACM* 39(3): 61-68.
- Bryant, D. J. (1992). "A spatial representation system in humans." *Psychology* 3(16).
- Bryant, D. J., Tversky, B. and Franklin, F. (1992). "Internal and external spatial frameworks for representing described scenes." *Journal of Memory and Language* 31: 74 - 98.
- Bryden, M. P. (1982). *Laterality: Functional asymmetry in the intact brain*. New York, NY, Academic.
- Burgoon, J., Buller, D. and Woodall, W. (1996). *Nonverbal communication: the unspoken dialogue*, McGraw-Hill Companies, Inc.
- Card, S., Mackinlay, J. and Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco, CA, Morgan Kaufmann.
- Caudell, T. P. and Mizell, D. W. (1992). Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes. In *Proceedings of International Conference on System Sciences*, 659-669. Kauai, Hawaii. January 1992.
- Cavell, R. (2002). *McLuhan in space*. Toronto, Canada, University of Toronto Press.
- Chance, S., Garnet, F., Beall, A. and Loomis, J. (1998). "Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration." *Presence: Teleoperators and Virtual Environments* 7: 168 - 178.

- Corballis, M. and Beale, I. (1983). *The ambivalent mind: The neuropsychology of left and right*. Chicago, Nelson-Hall.
- Corballis, M. C. (1993). *The lopsided ape: evolution of the generative mind*. New York, Oxford University Press.
- Cutting, J. E. and Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In (W. Epstein and S. Rogers eds.) *Perception of space and motion* (pp. 69-117). San Diego, CA, Academic Press. 5.
- D'Avossa, G. and Kersten, D. (1996). "Evidence in human subjects for independent coding of azimuth and elevation for direction of heading from optic flow." *Vision Research* 36: 2915 - 2924.
- Dictgans, J. and Brandt, T. (1978). Visual-vestibular interaction: effects on self-motion perception and postural control. In (R. Held and H. Leibowitz eds.) *Perception: Vol.8 Handbook of sensory physiology* (pp. 755 - 804). New York, NY, Springer-Verlag.
- Dimond, S. J. and Farrington, L. (1977). "Emotional response to films shown to the right or left hemisphere measured by heart rate." *Acta Psychologica* 41: 259.
- Ditchburn, R. and Ginsborg, B. (1952). "Vision with a stabilized retinal image." *Nature* 170: 35 - 37.
- Dolezal, H. (1982). "Living in a world transformed: perceptual and performatory adaptation to visual distortion."
- Easton, R. and Sholl, M. (1995). "Object-array structure, frames of reference, and retrieval of spatial knowledge." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21: 483 - 500.
- Ebenholtz, S. and Mayer, D. (1968). "Rate of adaptation under constant and varied optical tilt." *Perceptual and Motor Skills* 26: 507 - 509.
- Eilan, N., McCarthy, R. and Brewer, B. (1993). *Spatial representation*. Oxford, UK, Oxford University Press.
- Ellis, A., Young, A. and Anderson, C. (1988). "Modes of word recognition in the left and right cerebral hemispheres." *Brain and Language* 35: 254 - 273.
- Engen, T. (1982). *The perception of odors*. New York, NY, Academic Press.
- Farrell, M. and Robertson, I. (1998). "Mental rotation and the automatic updating of body-centered spatial relationships." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 227 - 233.

- Feiner, S., MacIntyre, B. and Seligmann, D. (1993). "Knowledge-based Augmented Reality." *Communications of the ACM* 36(7): 52-62.
- Feiner, S., MacIntyre, B., Tobias, H. and Webster, A. (1997). A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. In *Proceedings of International Symposium on Wearable Computers*, 208-217. Cambridge, MA. October 13-14, 1997.
- Ferguson, E. S. (1994). *Engineering in the mind's eye*. Cambridge, MA, MIT Press.
- Fisher, E., Haines, R. and Price, T. (1980). Cognitive issues in head-up displays. Moffett Field, NASA Ames Research Center.
- Fitts, P. (1954). "The information capacity of the human motor system in controlling the amplitude of movement." *Journal of Experimental Psychology* 47(6): 381 - 391.
- Fitts, P. and Peterson, J. (1964). "Information capacity of discrete motor responses." *Journal of Experimental Psychology* 67(2): 103 - 112.
- Foley, J. and McChesney, J. (1976). "The selective utilization of information in the optic array." *Psychological Research* 38: 251 - 265.
- Foxlin, E. (2002). Motion tracking requirements and technologies. In (K. Stanney eds.) *Handbook of virtual environments* (pp. 163 - 210). Hillsdale, NJ, Lawrence Erlbaum & Associates.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York, NY, Basic Books.
- Gazzaniga, M. and Sperry, R. (1965). "Language after section of the cerebral commissures." *Brain* 88: 237 - 294.
- Grabowska, A. and Nowicka, A. (1996). "Visual-spatial-frequency model of cerebral asymmetry: A critical survey of behavioral and electrophysiological studies." *Psychological Bulletin* 120: 434 - 449.
- Gray, H., Bannister, L., Berry, M. and Williams, P. (1995). *The anatomical basis of medicine and surgery*.
- Graziano, M. (1999). "Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position." *Proceedings of the National Academic of Sciences, USA* 96: 10418 - 10421.
- Graziano, M. and Gross, C. (1998). "Spatial maps for the control of movement." *Current Opinion in Neurobiology* 8(195 - 201).

- Graziano, M. S. and Gross, C. (1995). The representation of extrapersonal space: A possible role for bimodal visual-tactile neurons. In (M. Gazzaniga eds.) *The cognitive neurosciences* (pp. 1021-1034). Cambridge, MA, M.I.T. Press.
- Guiard, Y. (1987). "Asymmetric division of labour in human skilled bimanual action: the kinematic chain as a model." *Journal of Motor Behavior* 19(4): 486 - 517.
- Guiard, Y. and Ferrand, T. (1995). Assymetry in bimanual skills. In (D. Elliot and A. Roy eds.) *Manual Asymmetries in Motor Performance* (pp. 175 - 195). Boca Raton, FL, CRC Press.
- Haines, R., Fischer, E. and Price, T. (1980). Head-up transition behaviour of pilots with and without head-up display in simulated low-visibility approaches. Moffett Field, NASA Ames Research Center.
- Hall, E. (1963). "A system for the notation of proxemic behavior." *American Anthropologist* 65: 1003 - 1026.
- Hall, E. (1966). *The hidden dimension: man's use of space in public and private*. Garden City, NY, Doubleday.
- Hancock, P. and Meshkati, N. (1988). *Human mental workload*. New York, NY, North-Holland.
- Hari, R. and Jousmäki, V. (1996). "Preference of personal to extrapersonal space in a visuomotor task." *Journal of Cognitive Neuroscience* 8(3): 305 - 307.
- Harris, C. (1963). "Adaptation to displaced vision: Visual motor or proprioceptive change?" *Science* 140: 812 - 813.
- Hart, S. and Staveland, L. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In (P. Hancock and N. Meshkati eds.) *Human Mental Workload* (pp. 139 - 183). Amsterdam, The Netherlands, North-Holland.
- Hay, J. and Pick, H. (1966). "Visual and proprioceptive adaption to optical displacement of the visual stimulus." *Journal of Experimental Psychology* 71: 150 - 158.
- Hearn, D. and Baker, M. P. (1996). *Computer Graphics, C Version*, Prentice Hall.
- Hecaen, H. and Albert, M. (1978). *Human neuropsychology*. New York, NY, Wiley.
- Heckenmueller, E. (1965). "Stabilization of the retinal image; a review of method." *Psychological Bulletin* 63: 157 - 159.
- Heidegger, M. (1968). *What is a thing?* Chicago, IL, H. Regnery Co.

- Held, R. and Schrank, M. (1959). "Adaptation to disarranged eye-hand coordination in the distance-dimension." *American Journal of Psychology* 72: 603 - 605.
- Hoagland, H. (1933). "Quantitative aspects of cutaneous sensory adaptation." *Journal of General Physiology* 16: 911 - 923.
- Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In (K. Boff, L. T. Kaufman and J. Thomas eds.) *Handbook of Perception and Human Performance, Vol. 1* (pp.). New York, NY, Wiley.
- Hood, J. (1950). "Studies in auditory fatigue and adaptation." *Acta Otolaryngology, Supplement* 92: 1 - 57.
- Horvitz, E., Kadie, C., Paek, T. and Hovel, D. (2003). "Model of attention in computing and communication: from principles to applications." *Communications of the ACM* 46(3): 52 - 59.
- Inzuka, Y., Osumi, Y. and Shinkai, K. (1991). Visibility of head up display for automobile. In *Proceedings of 35th Annual Meeting of the Human Factors Society*.
- Johnson, A. and Proctor, R. (2004). *Attention: theory and practice*. Thousand Oaks, CA, Sage Publications.
- Kay, A. (1984). "Computer software." *Scientific America* 251(3): 40 - 47.
- Keijsers, N., Admiraal, M., Cools, A., Bloem, B. and Gielen, C. (2005). "Differential progression of proprioceptive and visual information processing deficits in Parkinson's disease." *European Journal of Neuroscience* 21(1): 239 - 248.
- Khan, A., Matejka, J., Fitzmaurice, G. and Kurtenbach, G. (2005). Spotlight: directing users' attention on large display. In *Proceedings of ACM CHI 2005*, 791 - 798. Portland, OR. April 2 - 7, 2005.
- Kirsh, D. (1995). "The intelligent use of space." *Artificial Intelligence* 73(1-2): 31 - 68.
- Kirsh, D. and Maglio, P. (1992). Some epistemic benefits of action: Tetris, a case study. In *Proceedings of Fourteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ.
- Kohler, I. (1964). "The formation and transformation of the perceptual world." *Psychological Issues* 3: 1 - 173.
- Kosslyn, S. M. (1987). "Seeing and imagining in the cerebral hemispheres: A computational approach." *Psychological Review* 94: 148-175.
- Krauskopf, J. and Riggs, L. (1959). "Interocular transfer in the disappearance of stabilized images." *American Journal of Psychology* 72: 248 - 252.

- Langolf, G. (1973). Human motor performance in precise microscopic work. Ann Arbor, MI, University of Michigan.
- Larish, I. and Wickens, C. (1991). "Attention and HUDs: flying in the dark?"
- Lehikoinen, J. (2000). Virtual pockets. In *Proceedings of Fourth International Symposium on Wearable Computers*, 165 - 170. Atlanta, GA. October 18 - 21, 2000.
- Leibowitz, H. and Post, R. (1982). The two modes of processing concept and some implications. In (J. Beck eds.) *Organization and representation in perception* (pp.). Mahwah, NY, Erlbaum.
- Levine, M., Jankovic, I. and Palij, M. (1982). "Principles of spatial problem solving." *Journal of Experimental Psychology: General* 111: 157 - 175.
- Mann, S. (2000). Telepointer: Hands-Free Completely Self Contained Wearable Visual Augmented Reality without Headwear and without any Infrastructural Reliance. In *Proceedings of Fourth International Symposium on Wearable Computers*, 177.
- Maravita, A. and Iriki, A. (2004). "Tools for the body (schema)." *Trends in Cognitive Science* 8(2): 79 - 86.
- Marks, L. (1978). *The unity of the senses*. New York, NY, Academic Press.
- Marks, L. and Armstrong, L. (1994). Haptic and visual representations of space. In (T. Inui and J. McClelland eds.) *Attention and Performance* (pp. 262 - 288). Cambridge, MA, The MIT Press.
- McCann, H. (1998). *The works of agency: on human action, will and freedom*. Ithaca, NY, Cornell University Press.
- McCann, R., Foyle, D. and Johnston, J. (1994). Attentional limitations with head-up displays. In *Proceedings of International Symposium on Aviation Psychology*. Columbus, OH.
- McCrickard, D. and Chewar, C. (2003). "Attentive user interface: attuning notification design to user goals and attention costs." *Communications of the ACM* 46(3): 67 - 72.
- McLuhan, M. (1967). *Gutenberg galaxy*. Toronto, Canada, University of Toronto Press.
- McNamara, T. (1986). "Mental representations of spatial relations." *Cognitive Psychology* 18: 87 - 121.
- McNamara, T. (1989). "Mental representations of spatial and nonspatial relations." *Quarterly Journal of Experimental Psychology* 41: 215 - 233.
- Melville, J. (1957). "Word-length as a factor in differential recognition." *American Journal of Psychology* 37: 85 - 106.

- Merickel, M. L. (1992). A study of the relationship between virtual reality (perceived realism) and the ability of children to create, manipulate and utilize mental image for spatially related problem solving. In *Proceedings of Annual Convention of the National School Boards Association*. Washington, DC.
- Mine, M. (1996). Working in a virtual world: interaction techniques used in Chapel Hill Immersive Modeling Program. Chapel Hill, NC, University of North Carolina, Chapel Hill.
- Mou, W., Biocca, F., Owen, C., Tang, A., Xiao, F. and Lim, L. (2004a). "Frames of reference in mobile augmented reality displays." *Journal of Experimental Psychology: Applied* 10(4): 238-244.
- Mou, W. and McNamara, T. (2002). "Intrinsic frames of reference in spatial memory." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 162 - 170.
- Mou, W., McNamara, T., Valiquette, C. and Rump, B. (2004b). "Allocentric and egocentric updating of spatial memory." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30: 142 - 157.
- Mou, W., Zhang, K. and McNamara, T. (2004c). "Frames of reference in spatial memories acquired from language." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30: 171 - 180.
- Mountcastle, V. (1976). "The world around us: neural command functions for selective attention." *Neuroscience Research Program Bulletin* 14: 1 - 47.
- Murphy, K. and Goodale, M. (1997). "Manual prehension is superior in the lower visual hemifield." *Society for Neuroscience Abstracts* 23: 178.
- Neisser, U. and Becklen, R. (1975). "Attention to visually specified events." *Cognitive Psychology* 7: 480 - 494.
- Neumann, U. and Majoros, A. (1998). Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. In *Proceedings of IEEE VRAIS '98*, 4 - 11. Atlanta, GA. March 14 - 18, 1998.
- Norman, D. (1993). *Things that make us smart: defending human attributes in the age of the machine*. Menlo Park, CA, Addison-Wesley Publishing Co.
- Osgood, C., Suci, G. and Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL, University of Illinois Press.
- Owen, C., Biocca, F., Tang, A., Xiao, F., Mou, W. and Lim, L. (2005). Information frames in mobile augmented reality user interfaces. In *Proceedings of Human Computer Interaction International 2005, 11th International Conference on Human-Computer Interaction*. Las Vegas, NV.

- Owen, C., Tang, A. and Xiao, F. (2003). ImageTclAR: a blended script and compiled code development system for augmented reality. In *Proceedings of STARS2003, The International Workshop on Software Technology for Augmented Reality Systems*, 23 - 28. Tokyo, Japan. October 7, 2003.
- Pani, J. and Dupree, D. (1994). "Spatial reference frames in the comprehension of rotational motion." *Perception* 23: 929 - 946.
- Pettigrew, J. and Dreher (1987). Parallel processing of binocular disparity in the cat's retinogeniculocortical pathways. In *Proceedings of Royal Society B: Biological Science*, 297 - 321. 22nd December, 1987.
- Pierce, J. and Pausch, R. (2002). Comparing Voodoo Dolls and HOMER: exploring the importance of feedback in virtual environments. In *Proceedings of ACM CHI 2002*. Minneapolis, MN. April 20 - 25, 2002.
- Pierce, J., Stearns, B. and Pausch, R. (1999). Voodoo Dolls: seamless interaction at multiple scales in virtual environments. In *Proceedings of ACM Symposium on Interactive 3D Graphics*.
- Poupyrev, I., Billinghamurst, M., Weghorst, S. and Ichikawa, T. (1996). The Go-Go Interaction Technique: non-linear mapping for direction manipulation in VR. In *Proceedings of the 9th annual ACM Symposium on User Interface Software and Technology*, 79 - 80. Seattle, WA.
- Presson, C. and Montello, D. (1994). "Updating after rotational and translational body movements: coordinate structure of perspective space." *Perception* 23(1447 - 1455).
- Previc, F. H. (1990a). "Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications." *Behavioral and Brain Sciences* 13: 519 - 542.
- Previc, F. H. (1990b). "Visual processing in three-dimensional space: perceptions and misperceptions." *Behavioral and Brain Sciences* 13: 559 - 566.
- Previc, F. H. (1998). "The neuropsychology of 3D space." *Psychological Bulletin* 124: 123 - 164.
- Previc, F. H. and Blume, J. (1993). "Visual search asymmetries in three-dimensional space." *Vision Research* 33: 2697 - 2704.
- Previc, F. H. and Neel, R. (1995). "The effects of visual surround eccentricity and size on manual and postural control." *Journal of Vestibular Research* 5: 399 - 404.
- Proctor, R. and Van Zandt, T. (1994a). *Human factors in simple and complex systems*. Boston, MA, Allyn and Bacon.

- Proctor, R. W. and Van Zandt, T. (1994b). Anthropometrics and workspace design. In eds.) *Human factors in simple and complex systems* (pp.). Boston, MA, Allyn and Bacon.
- Psotka, J. "Memory in VR and Augmented VR." from <http://alex-immersion.army.mil/serial.html>.
- Redelmeier, D. and Tibshirani, R. (1997). "Association between cellular telephone calls and motor vehicle collisions." *New England Journal of Medicine* 336(7): 453 - 458.
- Reeves, B. and Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge, UK, Cambridge University Press.
- Rieser, J. (1989). "Access to knowledge of spatial structure at novel points of observation." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: 1157 - 1165.
- Rieser, J. (1999). Dynamic spatial orientation and the coupling of representation and action. In (R. Golledge eds.) *Wayfinding behaviour: cognitive mapping and other spatial processes* (pp. 168 - 191). Baltimore, MD, Johns Hopkins University Press.
- Rieser, J., Guth, D. and Hill, E. (1986). "Sensitivity to perspective structure while walking without vision." *Perception* 15: 173 - 188.
- Rieser, J., Pick, H. and Ashmead, D. (1995). "Calibration of human locomotion and models of perceptual-motor organization." *Journal of Experimental Psychology: Human Perception and Performance* 21: 480 - 497.
- Riggs, L. and Ratliff, F. (1952). "The effects of counteracting the normal movements of the eye." *Journal of the Optical Society of America* 42: 872 - 873.
- Riggs, L., Ratliff, F., Cornsweet, J. and Cornsweet, T. (1953). "The disappearance of steadily fixed visual test objects." *Journal of the Optical Society of America* 43: 495 - 501.
- Rizzolatti, G. and Camarda, R. (1987). Neural circuits for spatial attention and unilateral neglect. In (M. Jeannerod eds.) *Neurophysiological and neuropsychological aspects of spatial neglect* (pp. 289 - 314). Amsterdam, The Netherlands, North-Holland.
- Rizzolatti, G., Gentilucci, M. and Matelli, M. (1985). Selective spatial attention: One center, one circuit, or many circuits? In (M. Posner and O. Marin eds.) *Attention and Performance II* (pp. 251 - 265). Hillsdale, NJ, Erlbaum.
- Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D. and van Dantzich, M. (1998). Data Mountain: Using spatial memory for document management. In *Proceedings of ACM UIST '98 Symposium on User Interface Software & Technology*. San Francisco, CA. November, 1998.

- Robertson, L. C. and Lamb, M. (1991). "Neuropsychological contributions to theories of part/whole organization." *Cognitive Psychology* 23: 299 - 330.
- Roel, V. (2002). Designing attentive interfaces. In *Proceedings of Symposium on Eye Tracking Research and Applications*. New Orleans, LA.
- Rolland, J., Biocca, F., Barlow, T. and Kancherla, A. (1995). Quantification of adaptation to virtual-eye location in see-thru head-mounted displays. In *Proceedings of Virtual Reality Annual International Symposium (VRAIS '95)*, 56-66. Research Triangle Park, NC. 11-15 March. IEEE Computer Society.
- Rubin, N., Nakayama, K. and Shapley, R. (1996). "Enhanced perception of illusory contours in the lower versus upper visual hemifields." *Science* 271: 651 - 653.
- Sackheim, H., Gur, R. and Saucy, M. (1978). "Emotions are expressed more intensely on the left side of the face." *Science* 202: 434 - 436.
- Schmalstieg, D. and Wagner, D. (2005). A handheld augmented reality museum guide. In *Proceedings of IADIS International Conference on Mobile Learning 2005*. Qawra, Malta. June 28 - 30, 2005.
- Sergent, J. (1983). "Role of the input in visual hemispheric asymmetries." *Psychological Bulletin* 93: 481-512.
- Sergent, J. (1987). "Failures to confirm the spatial-frequency hypothesis: Fatal blow or healthy complication?" *Canadian Journal of Psychology* 41: 412 - 428.
- Servos, P., Goodale, M. and Jakobson, L. (1992). "The role of binocular vision in prehension: a kinematic analysis." *Vision Research* 32: 1513 - 1521.
- Sheehan, J. and Sosna, M. (1991). *The boundaries of humanity: humans, animals, machines*. Berkeley, CA, University of California Press.
- Sheliga, B., Craighero, L., Riggio, L. and Rizzolatti, G. (1997). "Effects of spatial attention on directional manual and ocular responses." *Experimental Brain Research* 114: 339 - 351.
- Shelton, A. and McNamara, T. (2001a). "Systems of spatial reference in human memory." *Cognitive Psychology* 43: 274 - 310.
- Shelton, A. and McNamara, T. (2001b). "Visual memory from nonvisual experiences." *Psychological Science* 12: 343 - 347.
- Shiffrin, R. (1979). "Visual processing capacity and attentional control." *Journal of Experimental Psychology: Human Perception and Performance* 5: 522 - 526.
- Shneiderman, B. (1983). "Direct manipulation: A step beyond programming languages." *IEEE Computer* 16(8): 57 - 69.

- Shoemake, K. (1985). Animating rotation with quaternion curves. In *Proceedings of 12th Annual Conference on Computer Graphics and Interactive Techniques*, 245 - 254.
- Sholl, M. and Bartels, G. (2002). "The role of self-to-object updating in orientation-free performance on spatial memory tasks." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 422 - 436.
- Sholl, M. and Nolin, T. (1997). "Orientation specificity in representations of place." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23: 1494 - 1507.
- Simons, D. and Wang, R. (1998). "Perceiving real-world viewpoint changes." *Psychological Science* 9: 315 - 320.
- Sojourner, R. and Antin, J. (1990). "The effects of a simulated head-up display speedometer on perceptual task performance." *Human Factors* 32(3): 329 - 339.
- Solso, R. (1998). *Cognitive psychology*. Needham Height, MA, Allyn & Bacon.
- Sperry, R. (1961). "Cerebral organization and behavior." *Science* 133: 1749 - 1757.
- Stein, B. (1984). "Multimodal representation in the superior colliculus and optical tectum." *Journal of Neurophysiology* 41: 55 - 64.
- Stoakley, R., Conway, M. and Pausch, R. (1995). Virtual reality on a WIM: interaction worlds in miniature. In *Proceedings of ACM SIGCHI '95*.
- Stratton, G. (1897). "Upright vision and the retinal image." *Psychological Review* 4: 182 - 187.
- Strauss, E. (1998). "Writing speech separated in split brain." *Science* 280(5365): 827 - 828.
- Strayer, D. and Johnston, W. (2001). "Driven to distraction: dual-task studies of simulated driving and conversing on a cellular phone." *Psychological Science* 12(6): 462 - 466.
- Tang, A., Owen, C., Biocca, F. and Mou, W. (2003). Comparative Effectiveness of Augmented Reality in Object Assembly. In *Proceedings of ACM CHI '2003*, 73-80. Fort Lauderdale, FL.
- Telford, L. and Frost, B. (1993). "Factors affecting the onset and magnitude of linear vection." *Perception and Psychophysics* 53: 682 - 692.
- van der Heijden, A. (1992). *Selective attention in vision*. New York, NY, Routledge.
- van der Heijden, A. (2003). *Attention in vision perception, communication and action*. New York, N.Y., Psychology Press.

- Wada, Y., Saijo, M. and Kato, T. (1998). "Visual field anisotropy for perceiving shape from shading and shape from edges." *Interdisciplinary information sciences* 4(2).
- Wallach, H. (1987). "Perceiving a stable environment when one moves." *Annual Review of Psychology* 38: 1 - 27.
- Waller, D., Montello, D., Richardson, A. and Hegarty, M. (2002). "Orientation specificity and spatial updating." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 1051 - 1063.
- Wang, R. and Simons, D. (1999). "Active and passive scene recognition across views." *Cognition* 70: 191 - 210.
- Ware, C. (2000). *Information visualization*. San Francisco, Morgan Kaufmann.
- Ware, C., Arthur, K. and Booth, K. (1993). Fish tank virtual reality. In *Proceedings of SIGCHI conference on Human factors in computing systems*, 37-42. Amsterdam, The Netherlands. April 24-29, 1993.
- Weintraub, D. R., Haines, R. and Randle, R. (1985). Head-up display (HUD) utility II: Runway to HUD transitions monitoring eye focus and decision times. In *Proceedings of Human Factors Society 29th Annual Meeting*.
- Whitehead, R. (1991). "Right hemisphere superiority during sustained visual attention." *Journal of Cognitive Neuroscience* 3: 329 - 334.
- Wraga, M., Creem, S. and Proffitt, D. (2000). "Updating displays after imagined object and viewer rotations." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 151 - 168.
- Yamamoto, N. and Shelton, A. (2005). "Visual and proprioceptive representations in spatial memory." *Memory & Cognition* 33(1): 140 - 150.
- Yates, F. (1966). *The art of memory*. Chicago, IL, University of Chicago Press.
- Young, A. and Ellis, A. (1985). "Different methods of lexical access for words presented to the left and right visual hemifields." *Brain and Language* 24: 326 - 358.
- Yovel, G., Yovel, I. and Levy, J. (2001). "Hemispheric asymmetries for global and local visual perception effects of stimulus and task factors." *Journal of Experimental Psychology: Human Perception and Performance* 27: 1369 - 1385.
- Zhang, J. and Norman, D. (1994). "Representations in distributed cognitive tasks." *Cognitive Science* 18(1): 87 - 122.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02845 1510