

**DEVELOPMENT AND EVALUATION OF A NOVEL METHOD FOR *IN-SITU*
MEDICAL IMAGE DISPLAY**

by

Gaurav Shukla

B.S.E., Biomedical and Electrical Engineering, Duke University, 2004

Submitted to the Graduate Faculty of
Swanson School of Engineering in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH
SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Gaurav Shukla

It was defended on

March 16, 2011

and approved by

Howard Aizenstein, M.D., Ph.D., Associate Professor, Department of Bioengineering

Brian Chapman, Ph.D., Associate Professor, Division of Biomedical Informatics

Roberta Klatzky, Ph.D., Professor, Department of Psychology

Dissertation Director: George Stetten, M.D., Ph.D., Professor, Department of
Bioengineering

Copyright © by Gaurav Shukla

2011

**DEVELOPMENT AND EVALUATION OF A NOVEL METHOD FOR *IN-SITU*
MEDICAL IMAGE DISPLAY**

Gaurav Shukla, Ph.D.

University of Pittsburgh, 2011

Three-dimensional (3D) medical imaging, including computed tomography (CT) and magnetic resonance (MR), and other modalities, has become a standard of care for diagnosis of disease and guidance of interventional procedures. As the technology to acquire larger, more magnificent, and more informative medical images advances, so too must the technology to display, interact with, and interpret these data.

This dissertation concerns the development and evaluation of a novel method for interaction with 3D medical images called “grab-a-slice,” which is a movable, tracked stereo display. It is the latest in a series of displays developed in our laboratory that we describe as *in-situ*, meaning that the displayed image is embedded in a physical 3D coordinate system. As the display is moved through space, a continuously updated tomographic slice of a 3D medical image is shown on the screen, corresponding to the position and orientation of the display. The act of manipulating the display through a “virtual patient” preserves the perception of 3D anatomic relationships in a way that is not possible with conventional, fixed displays. The further addition of stereo display capabilities permits augmentation of the tomographic image data with out-of-plane structures using 3D graphical methods.

In this dissertation we describe the research and clinical motivations for such a device. We describe the technical development of grab-a-slice as well as psychophysical experiments to evaluate the hypothesized perceptual and cognitive benefits. We speculate on the advantages and

limitations of the grab-a-slice display and propose future directions for its use in psychophysical research, clinical settings, and image analysis.

TABLE OF CONTENTS

1.0	INTRODUCTION.....	1
1.1	THESIS STATEMENT.....	1
1.2	OVERVIEW OF CONTRIBUTIONS.....	2
1.3	THESIS ORGANIZATION	3
1.4	MATH NOTATION CONVENTIONS	4
2.0	BACKGROUND AND SIGNIFICANCE	5
2.1	RELATED WORK ON VISUALIZATION OF 3D DATA	8
2.1.1	Conventional computer graphics for 3D rendering.....	8
2.1.2	Stereo vision.....	9
2.2	RELATED WORK IN NAVIGATION THROUGH 3D DATA.....	11
2.3	CLINICAL SIGNIFICANCE.....	13
2.4	COMPUTERIZED IMAGE ANALYSIS.....	15
3.0	DEVICE DEVELOPMENT.....	18
3.1	THE SONIC PENLIGHT.....	18
3.2	CLINICAL TRIALS WITH THE SONIC FLASHLIGHT	20
3.3	VIRTUAL ULTRASOUND ENVIRONMENT	20
3.4	FIRST-GENERATION DEVICE	22
3.5	SECOND GENERATION DEVICE.....	25

3.6	COORDINATE FRAMES OF REFERENCE	29
3.7	STEREO DISPLAY AND 3D GRAPHICS COORDINATE SYSTEM.....	31
4.0	PSYCHOPHYSICS EXPERIMENT: EVALUATION OF VISUALIZATION AND NAVIGATION USING GRAB-A-SLICE	36
4.1	METHOD.....	36
4.1.1	Subjects	36
4.1.2	Stimuli	37
4.1.3	Design	38
4.1.4	Procedure.....	38
4.1.4.1	Navigation task.....	38
4.1.4.2	Visualization task.....	39
4.1.5	Training.....	39
4.1.6	Data Analysis	40
4.2	EXPERIMENTAL RESULTS	42
4.2.1	Navigation task.....	42
4.2.2	Visualization task	42
4.3	SUMMARY AND CONCLUSIONS OF EXPERIMENT	44
5.0	STEREO VISUALIZATION RESULTS.....	46
5.1	PRESERVATION OF PREVIOUS FUNCTION OF GRAB-A-SLICE	46
5.2	RENDERING OF 3D OBJECTS	48
6.0	DISCUSSION AND FUTURE DIRECTIONS.....	50
6.1	ADVANTAGES AND LIMITATIONS OF GRAB-A-SLICE	50
6.2	CLINICAL UTILITY	52

6.3 FUTURE PSYCHOPHYSICAL EXPERIMENTS	54
6.4 COMPUTER-AIDED SEGMENTATION ALGORITHMS	55
APPENDIX A	57
APPENDIX B	74
BIBLIOGRAPHY	75

LIST OF TABLES

Table 1. Coordinate frames of reference.....	29
--	----

LIST OF FIGURES

Figure 2.1: 3D perspective rendering (above) and sequence of CT images (below) of a human skull (skull.cg-graphic.com, answers.com).....	6
Figure 2.2: The Model 4 Sonic Flashlight (Wu, et al).....	7
Figure 2.3: Stereopsis for perception of depth due to retinal disparity (vision3d.com).	9
Figure 2.4: The red and blue objects fall on the horopter, forming corresponding retinal images; the brown object does not (sinauer.com).	10
Figure 2.5: Passive interface props.	12
Figure 2.6: The cubic mouse.....	12
Figure 2.7: Pulmonary embolus. At right, axial CT image showing embolus in the right pulmonary arterial tree (mdconsult.com, e-radiography.net).	14
Figure 3.1: The sonic penlight for guidance of superficial subdermal access (Shukla, et al).	19
Figure 3.2: First prototype of grab-a-slice, with 4DOF as labeled A-D.	23
Figure 3.3: User operating the grab-a-slice display, which can be manipulated freely in six degrees of freedom.....	27
Figure 3.4: The grab-a-slice display mounted onto its flexible boom arm. Inset: one hemisphere containing five IRED markers.	28
Figure 3.5: 3D shutterglasses with IRED markers rigidly mounted.	32
Figure 3.6: GL viewing frustum (blue) for 3D rendering in OpenGL.....	33
Figure 4.1 Sample subject data for visualization task using grab-a-slice (yellow squares) and conventional display (blue diamonds), showing subject response angles vs. true stimulus angles (as measured in degrees). Note the reversal error in the second quadrant.	41

Figure 4.2 Average response for each stimulus angle across all subjects, along with best-fit lines.
Solid line: grab-a-slice display $y = 0.95x - 0.58$. Dashed line: conventional display, $y = 0.86x - 1.51$ 43

Figure A.1: Visualization data for subject 1. 58

Figure A.2: Visualization data for subject 2. 59

Figure A.3: Visualization data for subject 3. 60

Figure A.4: Visualization data for subject 4. 61

Figure A.5: Visualization data for subject 5. 62

Figure A.6: Visualization data for subject 6. 63

Figure A.7: Visualization data for subject 7. 64

Figure A.8: Visualization data for subject 8. 65

Figure A.9: Visualization data for subject 9. 66

Figure A.10: Visualization data for subject 10. 67

Figure A.11: Visualization data for subject 11. 68

Figure A.12: Visualization data for subject 12. 69

Figure A.13: Visualization data for subject 13. 70

Figure A.14: Visualization data for subject 14. 71

Figure A.15: Visualization data for subject 15. 72

Figure A.16: Visualization data for subject 16. 73

PREFACE

During the nearly five years I have invested on my research training, I have honed some technical skills and learned some artistic approaches to engineering. None of that would have been possible without the guidance and help of many individuals.

First, I would like to thank my thesis advisor and mentor, George Stetten. His unwavering support and love of teaching have been vital to any success I have found. I am grateful to have had him as my advisor and even more glad to call him my friend. I am also grateful to the other members of my thesis committee: Howard Aizenstein, Brian Chapman, and particularly Roberta Klatzky. They have been invaluable to my work in their own way, and I thank them. Past and present members of our laboratory have each been extremely helpful to me in my work, especially John Galeotti, Damion Shelton, and Bing Wu.

Throughout my life, I have had the unconditional support of my parents, Mira and Rakesh, my brother Vaibhav, and my close friends to provide me with strength to face challenges and affirmation to celebrate my successes. This was never truer than during my thesis research, and for that I am grateful.

1.0 INTRODUCTION

1.1 THESIS STATEMENT

Display of 3D medical data on a movable tracked tomographic screen is more conducive to navigation through and visualization of those data than on a conventional immobile screen, and provides a novel platform for medical image analysis.

1.2 OVERVIEW OF CONTRIBUTIONS

This dissertation makes two major contributions. The first is the design, development, and validation of grab-a-slice, a movable tracked tomographic display that uses custom software and hardware to render slices of 3D data as a function of the position and orientation of the display. Through carefully designed psychophysical experiments, I demonstrated that using the grab-a-slice display confers a benefit to novices in their ability to accurately navigate through and visualize 3D structures as compared with a conventional display.

The second major contribution is the inclusion of stereo out-of plane graphics into grab-a-slice. To our knowledge, this combination of a movable tomographic display with stereo capabilities has never been done before. The stereo capabilities allow for the stable rendering of 3D objects in and out of the plane of the screen. This is a fertile testbed for future research on 3D medical image display and analysis.

1.3 THESIS ORGANIZATION

The next chapter describes the relevant background for the present research, including current and previous methods in navigation through and visualization of 3D data, clinical motivation, and relevant methods in image analysis. Chapter 3 describes the evolution of the design of the grab-a-slice display, and Chapter 4 encompasses the psychophysical experiment done to evaluate performance with the display. Chapter 5 describes the current results of the stereographical component of grab-a-slice. Chapter 6 offers discussion on the significance, strengths, and limitations of this work as well as insights into future directions.

1.4 MATH NOTATION CONVENTIONS

I have endeavored to abide by the following notation scheme throughout this document:

1. Matrices are identified with bold capital letters, e.g. \mathbf{T}
2. N-d vectors are identified with bold, italicized, lowercase letters, e.g. \mathbf{q}
3. N-d points are identified as vectors, as they are assumed to be vectors from the global origin to a particular location in space
4. A vector from \mathbf{a} to \mathbf{b} is computed by taking $\mathbf{b} - \mathbf{a}$
5. Points and vectors should be construed in either the homogeneous or non-homogeneous form, as appropriate for the context
6. Reference frames are identified by italicized capital letters, e.g. R
7. Points and vectors in a particular reference frame will have the reference frame as a subscript, e.g. \mathbf{a}_G
8. Transformations from reference frame A to frame B are denoted as $\mathbf{T}_{A \rightarrow B}$
9. Scalars are represented by non-bold lowercase letters, e.g. z

2.0 BACKGROUND AND SIGNIFICANCE

Three-dimensional (3D) medical imaging, including computed tomography (CT) and magnetic resonance (MR), and other modalities, has become a standard of care for diagnosis of disease and guidance of interventional procedures. However, effective visualization of 3D data remains limited by the available 3D display technology. Volumetric data can be rendered on two-dimensional (2D) displays by graphical techniques that make use of a number of cues for depth perception, including stereopsis, shading, opacity, and perspective, among others. Nonetheless, doctors generally prefer to view 3D data in tomographic slices primarily because graphical computer renderings may obscure relevant information from the viewer. For example, the entire brain is hidden by the skull in the rendering at the top of Figure 2.1. While tomographic viewing of 3D data is sufficient for many clinical tasks facing the radiologist, gaining proficiency with this method of interpretation requires overcoming a steep learning curve.



**Figure 2.1: 3D perspective rendering (above) and sequence of CT images (below) of a human skull (skull.cg-
graphic.com, answers.com).**

One particular use of tomographic viewing is with ultrasound (US), in which a slice is scanned and displayed in real-time, and in which practitioners exhibit the same learning curve mentioned above. Over the last decade, our laboratory has sought to reduce this learning curve by developing *in-situ* visualization methods for medical images that merge ultrasound images with a direct view of the patient [1][2]. Our device, called the Sonic Flashlight (SF), consists of a small display and a half-silvered mirror mounted on a conventional US transducer. Looking through the mirror, the operator sees the reflection of the real-time US image floating *in-situ* within the patient, precisely where the scan is currently being obtained (Figure 2.2). The SF merges the US image, US probe, operator's hands, surgical instrument, and patient into the same field of view, enabling perceptually guided action. Others have taken similar approaches for displaying slices of CT and MR data *in-situ* [3][4][5].

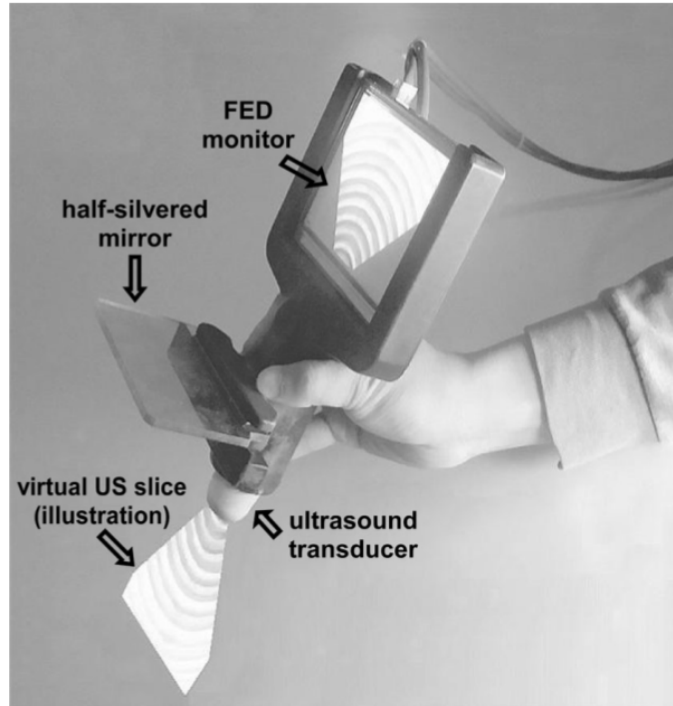


Figure 2.2: The Model 4 Sonic Flashlight (Wu, et al).

Our laboratory has conducted extensive research into the underlying psychophysical properties of *in-situ* image guidance. This research has demonstrated advantages as compared to conventional displays in the accuracy of perceived target depth, immunity to errors due to surface deformation, and the interpretation of shape and pose of 3D targets [6][7][8][9][10][11]. We use the term “tomographic aperture” to denote the manner in which 3D data is sampled by slicing, analogous to that of a conventional aperture through which the world is sampled by constrained projection. The research has established that displaying an image *in-situ* provides a spatial buffer for memory to combine sequential 2D information into a 3D context.

This avenue of research led us to conceive of a novel type of *in-situ* display that may provide a better interface for viewing and interacting with pre-acquired 3D data, such as those generated by CT or MR. We call this display device “grab-a-slice,” and it consists of a tracked mobile flat-panel monitor that displays slices through a stationary 3D data set. Grab-a-slice creates an easily manipulable tomographic aperture by means of a real image, as opposed to the

SF's virtual image, and thus produces an *in-situ* image with larger size and greater clarity (resolution, brightness, and contrast) than possible with the SF.

Conventionally, clinicians view tomographic images on an immobile 2D display, scrolling through stacks of parallel slices. However, we believe this method to be suboptimal for certain clinical tasks, such as locating and interpreting the boundaries of irregularly-shaped structures or following curved vessels along their trajectories. Furthermore, conventional displays do not preserve the context in which the data exists; users lose track of the location of relevant 3D structures in adjacent slices.

Grab-a-slice has evolved to provide an improved interface for many common clinical imaging tasks, as it allows the user to navigate more intuitively through 3D data. Grab-a-slice may also be a medium through which graphical augmentation can illuminate the information content of 3D medical images rather than obscure it. For example, grab-a-slice can be used to display the results of 3D segmentation algorithms in a novel fashion, potentially creating new (two-way) modes of interaction between users and image data. Thus, grab-a-slice may be useful for supervising semi-automated segmentation of relevant anatomy, to facilitate computer-aided diagnosis and therapy planning.

2.1 RELATED WORK ON VISUALIZATION OF 3D DATA

2.1.1 Conventional computer graphics for 3D rendering

Various methods have been developed for rendering 3D data onto a stationary 2D display, with or without special hardware for stereovision, such as Lorensen's marching cubes algorithm for

pre-segmented 3D objects [12] and Levoy's conversion of intensity gradient to rendered opacity and surface normal [13]. Many other graphical techniques may be applied to rendering medical data, but in general the data must first be analyzed to segment the structures to be displayed. After all, without some form of at least partial opacity by which objects can obscure other objects, the world would be a confusing cloud of translucent visual information very difficult for the human to interpret. As a result of the requirement to establish the locations of these opaque surfaces within the voxel data, graphical methods of 3D renderings are not widely used by clinicians, who generally prefer to see the raw data as slices than to trust a computer algorithm to segment and render structures correctly.

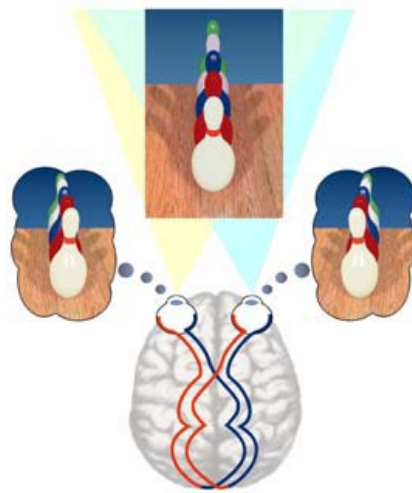


Figure 2.3: Stereopsis for perception of depth due to retinal disparity (vision3d.com).

2.1.2 Stereo vision

With grab-a-slice, we use computer graphics not as a primary method of rendering data, but rather to enhance the spatial and anatomic context of the *in-situ* tomographic image on the display. To do this, we make use of stereo display technology. One of the most powerful and easily manipulated cues for depth perception is stereopsis, in which humans use the binocular

disparity between two 2D retinal images to generate an impression of three-dimensionality in the visual world (Figure 2.3). When focusing on an object at a particular distance, there is a locus of points, called the *horopter*, for which the rays project to corresponding retinal points (that is, the points on the horopter are seen without stereo discrepancy). Within a certain range of the horopter, disparity between the retinal images causes the perception of depth relative to the horopter (Figure 2.4). Within this region, called *Panum's fusional area*, the discrepancies between retinal images are relatively small, and objects appear as fused, even though the retinal images do not perfectly correspond. Outside Panum's area, double vision, or *diplopia*, results, and depth relative to the horopter is determined by a higher-level process [14].

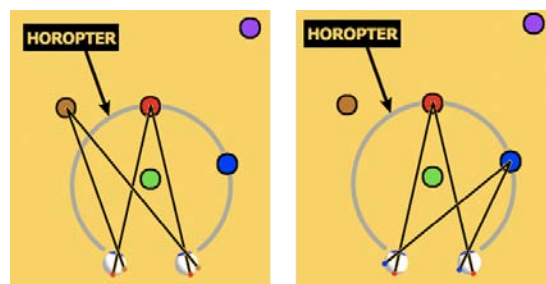


Figure 2.4: The red and blue objects fall on the horopter, forming corresponding retinal images; the brown object does not (sinauer.com).

Computer displays can present stereo images in several ways. Each method makes use of the human ability to interpret binocular disparity as the perception of depth. Different images can be presented to each retina by the use of specialized glasses, which filter the rays entering each eye by wavelength or polarization. An example of the latter is LCD shutterglasses (e.g., CrystalEyes, Inc., Beverly Hills, CA), which alternately darken the lens over each eye as the display alternates between two images. Commercial systems for consumer use are now widely available, with a large niche in the computer video game market.

Blackwell et al [15] designed a display system in which stereo glasses were utilized to overlay 3D medical image information into a simulated patient using a half-silvered mirror. We

extend this concept to grab-a-slice, in which the tracked movable display can include stereo images of objects extending outside the slice, augmenting the tomographic content in the slice itself.

2.2 RELATED WORK IN NAVIGATION THROUGH 3D DATA

Navigation through a displayed 3D environment has also inspired various approaches. Ware and Osborne [16] provide a user interface for exploring virtual graphics environments they call “scene-in-hand,” a virtual camera control that changes the perspective of a 3D environment in response to the manipulation of a tracked tool. Hinckley, et al., use passive interface props [17], tracked objects that are simple and hand-held, to generate tomographic slices of medical image data (Figure 5). The cubic mouse [18], developed for specification of 3D coordinates in graphics applications, consists of a tracked in-hand device coupled with rods and buttons to specify motion of virtual objects along various axes (Figure 6). Other approaches, including the SpaceBall line of products (3DConnexion, Silicon Valley, CA) implement a non-tracked 3D navigation tool. All of the above still use a stationary display, as opposed to the movable display described here.



Figure 2.5: Passive interface props.

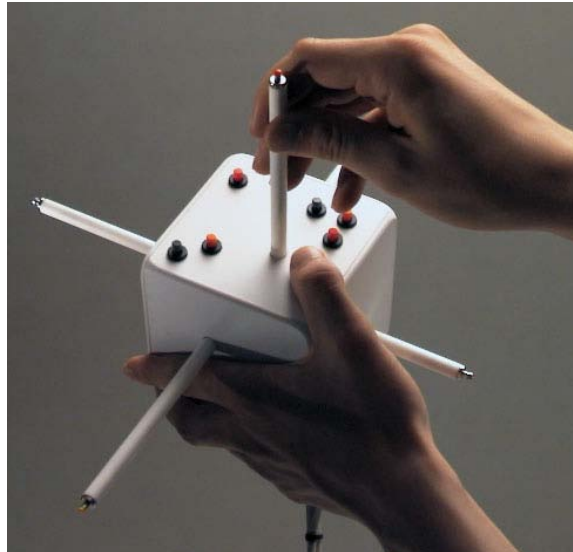


Figure 2.6: The cubic mouse.

Tracked movable boom-mounted monitors that are counterbalanced so that the operator can manipulate them by hand have been used as immersive displays into 3D virtual environments [19][20]. However, to our knowledge, they have not been used for tomographic slicing of a volume.

Augmented reality systems in which head-mounted displays are coupled with algorithms for 3D perspective rendering have been studied extensively [21][22]. This approach has also been applied to viewing tomographic slices [23][24].

Ito et al developed a system that tracks a handheld screen surface using a monocular vision system and then projects the corresponding slice of a 3D virtual object onto that screen in real-time. However, this system does not provide optimal image contrast and brightness, and has a limited region of operation [25].

Grab-a-slice differs from previous work in this field. To our knowledge, a movable tracked active display with stereo capability has not been used to show the unadulterated details of the medical image data in tomographic slices while simultaneously using 3D graphics to provide contextual information about the objects around the current slice.

2.3 CLINICAL SIGNIFICANCE

Pulmonary embolism (PE) is a condition in which an embolus (made of clotted blood, air, or fat) propagates into the pulmonary arterial tree, causing a partial or complete occlusion of blood flow to the lung (Figure 2.7). Depending on the size and location of the embolus, the condition ranges from asymptomatic to fatal. In a study of more than 42 million deaths (from all causes) that occurred over a 20-year duration, almost 600,000 patients (approximately 1.5 percent) were diagnosed with PE. PE was the presumed cause of death in approximately 200,000. This study underestimates the true incidence and prevalence of PE, since more than half of all PE are probably undiagnosed [26]. The average annual incidence in the US is approximately 1 in 1000, with about 250,000 incident cases occurring annually. Patients being evaluated for PE are typically given a contrast-enhanced CT scan, in which the blood vessels are filled with a high-contrast agent that makes them easily distinguishable in the CT image.

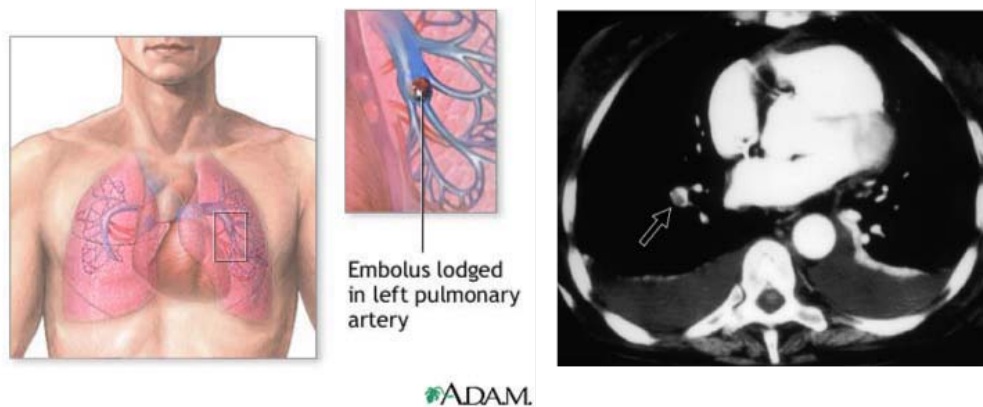


Figure 2.7: Pulmonary embolus. At right, axial CT image showing embolus in the right pulmonary arterial tree (mdconsult.com, e-radiography.net).

Radiologists follow a specific algorithm for detecting PE in the CT image in which they first identify suspicious features in the branches of the pulmonary vessels. PE can manifest in many ways, but is commonly observed as a filling defect caused by the presence of the embolus. The radiologist must trace the vessel containing the defect back to the heart to determine whether it is in fact a branch of the pulmonary artery (PA), emerging from the right ventricle, or of a pulmonary vein (PV), supplying the left atrium. PE is only seen in the branches of the PA, so suspicious structures in branches of the PVs are likely to be artifactual or otherwise non-significant.

Clinicians typically perform the task of vessel identification by scrolling through a stack of axial slices in the CT dataset. The clinician must invoke at least two cognitive processes: application of top-down knowledge of the vasculature and surrounding anatomy, and reconstruction of the path of 3D vessels through multiple 2D cross-sections. A hypothesis of the present thesis is that grab-a-slice can reduce this cognitive load by allowing the user to move in actual 3D space through the image volume.

Radiological diagnosis of PE is but one of many applications for which grab-a-slice may be well suited. Any medical imaging task requiring the reconstruction of 3D anatomy from 2D slices may lend itself to this device.

2.4 COMPUTERIZED IMAGE ANALYSIS

As mentioned above, we intend to adapt grab-a-slice as a tool for supervised image analysis. Segmentation algorithms hold great clinical promise in assisting radiologists with identifying normal and pathological anatomy, but currently most algorithms fall short of fulfilling this promise, especially when abnormal anatomy is involved. Various classes of segmentation methods, based on thresholding, region growing, maximum likelihood classifiers, Markov random field models, deformable models, atlas-based methods, and other techniques, have been developed in research settings but have generally yet to be sufficiently validated or widely accepted in the clinic [30][31].

Current attempts at segmentation of the vasculature have concentrated on methods that take advantage of the known tubular structure. Multiscale filtering is a method in which images are convolved with 3D Gaussian filters at multiple scales. Eigenvalues of the resulting Hessian matrix at each voxel are used to provide connectiveness along the axes of tubular structures [32][33]. Tobogganing is a segmentation algorithm in which the lowest gradient magnitude in each local 8-connected neighborhood of pixels determines a *slide direction*, connecting sets of pixels to avoid the boundaries [34][35][36].

Our laboratory has developed a method of semi-automated real-time segmentation of the common carotid artery and internal jugular vein in B-mode ultrasound images. This method,

known as the *spokes ellipse algorithm*, draws radial lines, or spokes, emanating from a seed point in the cross-sectional image of a vessel to detect an intensity boundary along each spoke. These boundary points are used to fit an ellipse, the center of which then acts as the seed point for the subsequent timeframe [37][38]. The method was used for temporal tracking of large vessels in 2D US images, but could be applied as well in segmentation of the pulmonary vessels in spatially separated slices. The above methods have practical limitations including extensive parameterization and difficulty with branching structures.

Our laboratory has also developed a method, Shells and Spheres, for extracting anatomical shape models from 3D images, which is especially well suited for the heart and vasculature. The method represents 3D shapes as the union of spherical regions, and is therefore well suited for tubular structures that can branch or curve back upon themselves. These spheres are found using a set of spherical operators centered at each image pixel and sized to reach, but not cross, the nearest object boundary. Statistical tests are performed on the sample of pixels contained by each sphere to measure homogeneity within objects and the heterogeneity between objects.

One particular Shells and Spheres algorithm developed in our laboratory by Aaron Cois [39][40][41][42] and then extended by Vikas Shivaprabhu [43] and others [44] is briefly described here. The basic algorithm may be outlined as follows: Spheres are “grown” at each pixel location by incorporating “shells” of pixel intensity values into expanding spheres, including an ongoing analysis of mean, variance, and first-order moment, to identify the radius at which each sphere grows past its nearest object boundary. Pairs of spheres on opposite sides of putative boundaries (i.e., “inner” and “outer” spheres) are then analyzed to estimate the likelihood of a local boundary existing; two spheres whose pixel populations differ in intensity

will have a larger likelihood of lying across a boundary from each other. Abrupt changes in the direction vectors between large-likelihood sphere pairs are then used to generate the set of points that constitute the *medial ridge* of the object, which identifies the interior of anatomical structures of interest. The union of spheres whose centers lie on the medial ridge and are tangent to the boundary in two or more locations can be used to represent the entire object, and this set of spheres can be visualized with 3D graphics.

The amount of human involvement in a segmentation algorithm ranges from none (in fully automated algorithms) to a great deal (fully manual systems). Shells and Spheres presents a number of promising features for operation with grab-a-slice. It is inherently multidimensional, permitting extension from the current slice into the surrounding 3D volume independent of the current slice's orientation. It is also very efficient, provided it can be limited to particular structures and regions, which grab-a-slice can be used to do during initiation and supervision.

Grab-a-slice may provide new ways for human input to be captured and utilized by segmentation routines that have not yet been considered in current approaches. 3D segmentations of structures can be displayed with stereo graphics on grab-a-slice. These segmentations can be overlaid with appropriate opacity and blending onto tomographic data to provide context-rich information that facilitates user interpretation of the image.

3.0 DEVICE DEVELOPMENT

As mentioned above, grab-a-slice is the latest in a series of devices our lab has developed that utilize *in-situ* images. This chapter describes the evolution of the current generation of the device, including relevant previous devices in whose development and validation I was involved, and which helped motivate grab-a-slice. This chapter then provides details of the current design.

3.1 THE SONIC PENLIGHT

The sonic flashlight has been described extensively (see Chapter 2.0). The clinical SF was designed to provide image guidance for particular clinical applications (e.g., vascular access and amniocentesis) that are performed at a particular scale – generally, between two and six centimeters below the patient’s skin. However, many clinical procedures require intervention within the first subdermal centimeter. Gaining vascular access in premature neonates, for example, is impeded by their vessels’ small size and mobility. Ophthalmologists are exploring real-time ultrasonic guidance in corneal surgery and in accessing structures in the anterior chamber of the eye [45]. Surgery for carpal tunnel syndrome involves tendons and nerves that reside in the first subdermal centimeter [46].

These examples and others may benefit from *in-situ* image guidance as provided by the sonic flashlight, but at a smaller scale. The first models of the SF [1][47] employed medium-

range ultrasound frequencies, which provide imaging information at a greater depth at the cost of poorer spatial resolution. Obtaining meaningful images at superficial depths with these models is thus not optimal. The footprint of the transducers is also not optimal for getting access just beneath the skin.

The need for image guidance techniques for superficial interventions motivated research toward construction of a miniature sonic flashlight that illuminates low-depth areas with high resolution. I worked with others in our laboratory to construct a working model of this device using a 12x9mm organic LED microdisplay, a custom CAD/CAM designed housing, and a small phased array ultrasound probe. With this device, we guided insertion of a needle into a gel phantom blood vessel (diameter = 6mm) (Figure 3.1) [48].

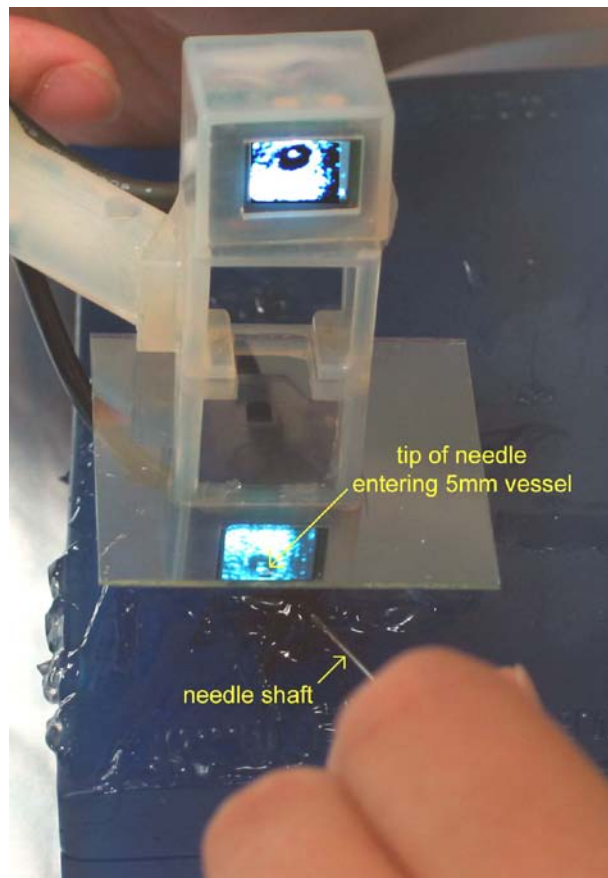


Figure 3.1: The sonic penlight for guidance of superficial subdermal access (Shukla, et al).

Experience with the sonic penlight led to a current project in our laboratory that replaces ultrasound with optical coherence tomography as the imaging modality. The device is designed to be used to image superficial ocular structures in corneal surgery [49].

3.2 CLINICAL TRIALS WITH THE SONIC FLASHLIGHT

In addition to the design and testing of the sonic penlight, I was involved in the clinical trials of the SF. The first clinical trials of this device were performed in the interventional radiology suite of Presbyterian Hospital at the University of Pittsburgh. Others in our laboratory had demonstrated the initial safety and feasibility of using the SF to guide the placement of peripherally inserted central catheters (PICCs) [50]. I was involved in a subsequent redesign and calibration of the SF and in trials of PICC placement in the IR suite [51] and at the patient's bedside by nurses in over 50 patients [52]. I was also involved in the first live-animal study using the SF for real-time guidance of an intervention on the kidney, in which we gained percutaneous access to the renal collecting system in a pig model [53]. My participation in these trials led to valuable experience with how operators interact with *in-situ* images in the clinical environment.

3.3 VIRTUAL ULTRASOUND ENVIRONMENT

While the above research gave me useful experience with the *in-situ* display of real images, grab-a-slice evolved most directly from another augmented reality system developed in our laboratory by Damion Shelton, the “virtual sonic flashlight.” The virtual SF differs from the clinical device

in that simulated data is examined in lieu of real ultrasound data in order to more easily study the psychophysics of *in-situ* image displays and how various image parameters affect a user's perception of the contents of a simulated ultrasound image [54]. The virtual SF utilizes a cluster of infrared light-emitting diodes (IREDs) and an optical tracking system to continuously monitor the position and orientation of a dummy ultrasound probe; the contents of the virtual SF's display are based upon this position and orientation data such that when the user is operating the device. Simulated targets were placed into a virtual patient, and using the virtual SF system, the operator can see registered simulated ultrasound images of cross sections of those targets below the scanning surface.

The virtual SF was used in a set of psychophysical experiments that illuminated some of the differences between *in-situ* visualization (using the SF) and *ex-situ* visualization (using a conventional display) [54]. In particular, one experiment attempted to determine the effect of background cues in the virtual image on the accuracy of hitting a virtual target with a needle tool. The effect was studied using targets at different depths in the virtual patient, comparing performance between the SF and conventional ultrasound. The initial results indicated, although not with statistical significance, that performance with SF guidance is more accurate than that with conventional ultrasound guidance, provided that rich background cues are present in the virtual image. The magnitude of this difference increases with target depth.

I became involved in this project in order to further validate the initial findings. I tested 22 subjects, of which 16 provided suitable data for statistical analysis. The results of this analysis supported the trends found by Shelton, though not to the desired level of significance. Nonetheless, I gained valuable experience in planning and performing psychophysical experiments on human subjects, as well as with the apparatus, which inspired the design of my

device as well as of my own experiments. Many of the basic features of the engineering, including geometric transformations, optical tracking, graphical programming, 3D texture mapping, etc., carried over into my dissertation research.

3.4 FIRST-GENERATION DEVICE

The first prototype of grab-a-slice [55] was constructed using a wooden frame for a 15” touch-screen display (Microtouch M150, 3M, Inc., St. Paul, MN) that allowed the display to be manipulated in 4 degrees of freedom (DOF) with respect to a stationary tabletop. The operator could move the display with a single hand using one or both of two handles (Figure 3.2). The apparatus was free to translate across the tabletop in two directions (A and B), and to rotate about the “yaw” axis (C), facilitated by Teflon pads attached to the underside of the platform to provide low dynamic friction. This permitted comfortable one-handed manipulation. A hinged mount with ball-bearing tracks permitted the display to rotate vertically about the “pitch” axis (D). Sufficient static friction in the hinge and between the platform and the tabletop, along with proper balance of the hinged assembly, permitted the screen to remain immobile when released. Rigid bodies can have a maximum of six DOF, but for the first prototype we deemed it unnecessary to physically implement the remaining two: translation in height normal to the tabletop and rotation in the “roll” direction within the image plane, as both could be readily implemented in software. A seventh DOF, isotropic scale, could also be manipulated in software, effectively magnifying or shrinking the virtual patient.



Figure 3.2: First prototype of grab-a-slice, with 4DOF as labeled A-D.

The first prototype of grab-a-slice used four IRED markers, roughly corresponding to each corner of the display, for detection by an optical tracking system (Optotrak Certus, Northern Digital Inc.). A subsequent revision increased the number of IRED markers to 10, divided between two hemispheres, providing greater range-of-motion and accuracy when manipulating and tracking the display. The Optotrak localized each IRED marker with an accuracy of approximately 0.1mm and a sampling frequency of at least 100 Hz, using a rigid array of three cameras fixed with respect to the volume of space through which grab-a-slice would be manipulated. All the markers were mounted on the portion of the grab-a-slice apparatus that is rigidly attached to the display itself. Thus the markers could be treated as a rigid body by the

Optotrak software to compute orientation and position for the display as a whole relative to the camera array.

An optically tracked stylus was used to find the four corners of the display. The stylus was tracked along with the 10-marker array so that the position of the display corners could be identified relative to the marker array. For the original prototype, the display was assigned to be at the mid-axial slice across the thorax when placed at the center of the table with zero pitch and zero yaw. The long-axis of the patient was assigned to the “range” or z-axis of the camera coordinate system. As the apparatus was moved to other locations and orientations, the Optotrak software continually reported the locations of the four corners of the display based on their relationships to the array of IRED markers. To display the appropriate slice from the 3D data on the display, the locations of the four corners of the display were used to extract the slice from the 3D data set by means of a 3D texture mapping board (GeForce 8800, NVIDIA, Inc., Santa Clara, CA). The method of 3D texture mapping interpolates voxels from a 3D data set onto polygons in arbitrary planes for 2D display, in this case a single rectangle occupying the surface of the screen [56]. Furthermore, the scale of the displayed data was adjusted accordingly so that the data would be shown as life-size (the scale could also be altered intentionally).

Using the commercial software provided by Northern Digital, a “rigid body file” was created for the grab-a-slice display that described the physical arrangement of the markers relative to one another. The Optotrak system uses this file when the software is initialized to report not only the position of the markers, but also a 4x4 homogeneous transform matrix describing the position and orientation of the rigid body as a whole. A 4x4 homogeneous transformation matrix \mathbf{T} can be decomposed into a 3x3 rotation component, \mathbf{R} , and a 3x1 translation column vector, \mathbf{t} .

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix}$$

One may think of this matrix as a transform between the camera's internal coordinate reference frame C to a reference frame W whose origin is within the rigid body (in this case, at the physical average of the array of markers describing the rigid body), $\mathbf{T}_{C \rightarrow W}$. The various reference frames developed for this work are described in more detail below.

3.5 SECOND GENERATION DEVICE

The second prototype of grab-a-slice replaced the tabletop wooden frame with a flexible boom arm (Figure 3.3). This change was important for subsequent psychophysical experiments for a number of reasons. Recall that we seek to implement the effect that the display is slicing through an "invisible, virtual patient." In order to capitalize on the potential perceptual benefits of *in-situ* imaging, the user must believably immerse himself into the reference frame of this invisible patient. However, the original tabletop prototype created several alternative frames of reference that might actually conflict with the frame of reference created by the data, making it difficult to cognitively reference the correct reference frame. Competing coordinate systems cause direct conflict in thinking about the correct one. The large bezel of the monitor and the wooden frame were natural focus points for the user that prevented intuitive immersion into the reference frame of the 3D data; rather, the user's gaze was often fixed in the coordinate frame of the display. In addition, the bulkiness of the frame and weight of the display made the device relatively difficult to move, despite the use of Teflon pads to reduce friction. The wood-frame prototype was set up in a busy lab space with many additional objects in the user's peripheral

field of view, including the tabletop and other furniture, which created even more alternate reference frames that confused the user.

Given these limitations, the hardware design of grab-a-slice was improved by mounting the display on a flexible, counterbalanced boom arm. As opposed to the wooden frame, which presented a large additional visual target to the user, the boom arm attached from the rear of the screen and was largely invisible to the user. The boom arm also expanded the motion of the screen to include all six degrees of freedom. This enabled comfortable one-handed or two-handed manipulation of the display through a large volume of space. The boom arm did, however, introduce irregular forces that depended on the position of the screen and the orientation of the boom, although these forces are not overwhelmingly intrusive during normal manipulation of the device.

Software improvements to the device were also critical in improving the immersion into the frame of the virtual patient. A precise calibration using coordinate transform matrices ensured that the location of the virtual patient was stable – the original approximation resulted in “patient drift” as the device was manipulated. I manipulated the contrast windows of the CT data to be optimal for viewing pulmonary vasculature. These changes along with the hardware improvements made it easier for users to find and remain in the correct reference frame.

The new prototype was moved to a dedicated laboratory space at Carnegie Mellon University (CMU). This location facilitated recruitment of experimental subjects and provided better control for experiments. The space in which grab-a-slice was manipulated at CMU was relatively devoid of objects that could present alternate reference frames to the user. The darkened room facilitated immersion into the 3D data space itself as presented on the screen.



Figure 3.3: User operating the grab-a-slice display, which can be manipulated freely in six degrees of freedom.

The touchscreen display from the first prototype was also replaced for the second prototype with a 120Hz 22" stereo-capable LCD (Samsung Syncmaster 2233rz, Ridgefield Park, NJ). In addition to being higher quality and larger, the higher speed of the new screen enabled stereo applications that will be described below. The pair of Styrofoam hemispheres containing IRED markers was rigidly mounted onto this display (Figure 3.4).

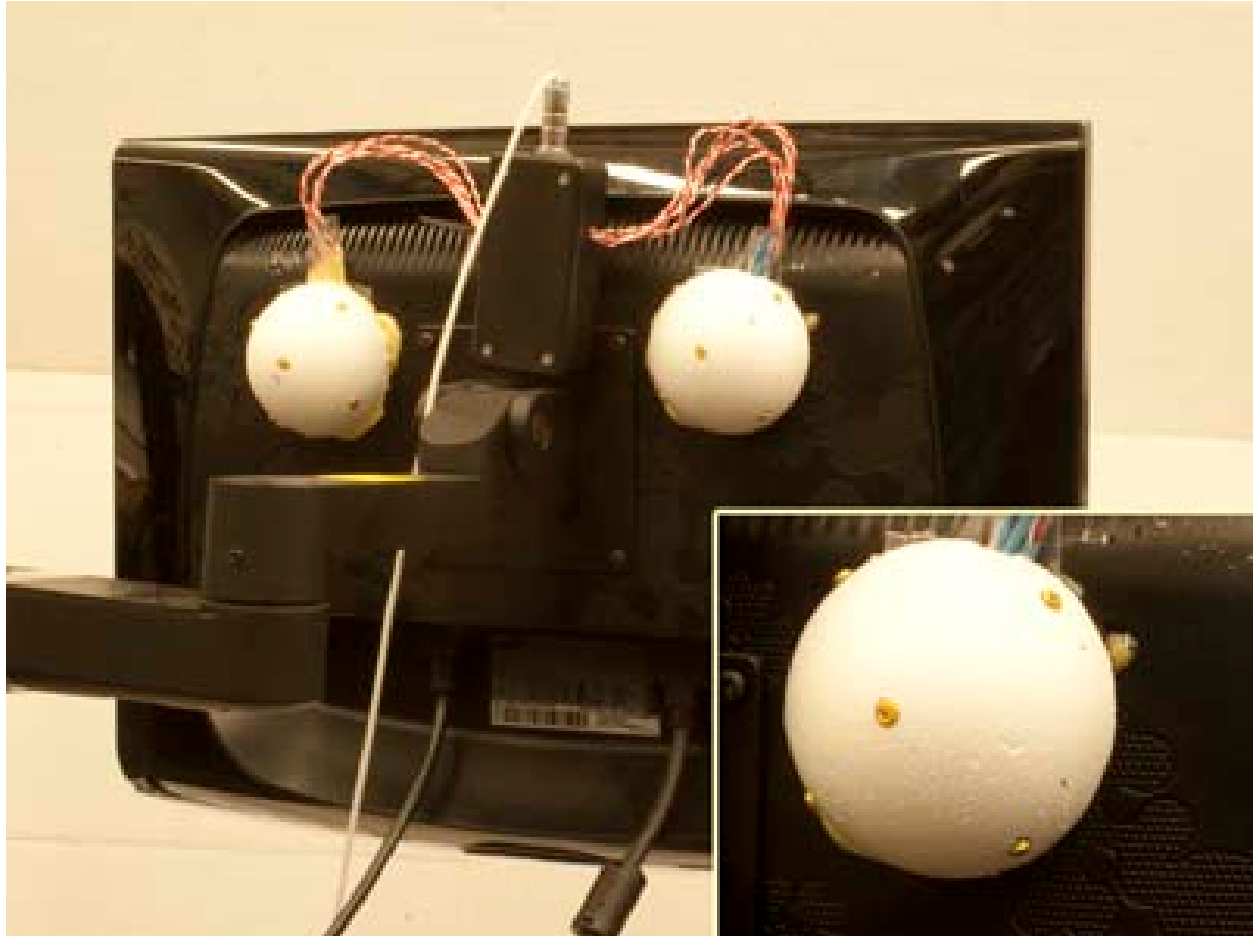


Figure 3.4: The grab-a-slice display mounted onto its flexible boom arm. Inset: one hemisphere containing five IRED markers.

To permit stereo applications, I also replaced the GeForce graphics card with a Quadro FX 3800 (NVIDIA, Inc.), which was compatible with the NVIDIA 3DVision emitter and shutterglasses while also performing the texture mapping described above.

3.6 COORDINATE FRAMES OF REFERENCE

Table 1. Coordinate frames of reference.

C	Camera coordinate system	mm
W	World coordinate system	mm
D	Data coordinate system	Integer voxel indices
G	Graphics coordinate system	mm
T	Texture memory coordinate system	Floating point $\in (0,1)$

A number of coordinate reference frames were required, as summarized in Table 1, and discussed below, to describe the position of tracked objects (the grab-a-slice display, and later, the 3D shutterglasses), as well as the individual voxels in the image and the graphical overlays in 2D and 3D. The set of coordinate frames of reference evolved from the first prototype of grab-a-slice to the current version.

The Optotrak Certus camera array has a built-in coordinate system, which I denote as C , with its origin at the lens of the center camera and units in millimeters. Using commercial software provided by Northern Digital Inc., I first characterized grab-a-slice in C , denoting the four corners of the display as $q_{i,C}$ for $i = 1,2,3,4$, corresponding to the lower left, lower right, upper left, and upper right, respectively. Because this coordinate system is based upon the position of the camera relative to the operating volume, the key values are not intuitive to the user.

I created a second coordinate reference frame called “world coordinates,” W , which is used to define the region of space through which the grab-a-slice display is to be manipulated. This allows the 3D image to be located conveniently for the user. Each time a new dataset is

loaded, the origin of W is reset to the lower-left corner of the grab-a-slice display at its current position: $\mathbf{T}_{C \rightarrow W}(\mathbf{q}_{1,C}) = \mathbf{q}_{1,W} = (0,0,0)_W$. Typically, this position is set at a convenient starting point for user interaction with image data. The other three corners $\mathbf{q}_{i,W}$ for $i = 2,3,4$ are calculated using the physical dimensions of the screen in millimeters. To compute the transform matrix $\mathbf{T}_{C \rightarrow W}$, four non-coplanar points are required. We use three corners \mathbf{q}_i for $i = 1,2,3$ and generate a fourth point outside of that plane by means of the cross product of two vectors defined by the first three points. These four corresponding points in C and W are used to solve for the transformation matrix \mathbf{T} by matrix inversion. When a new dataset is loaded, $\mathbf{T}_{C \rightarrow W}$ is recalculated given the user's starting position.

The coordinate reference frame for addressing particular voxels in the 3D image is referred to as D , the data coordinate system. In general, 3D medical images are stored with integer indices addressing each particular voxel. Radiologists have developed a convention for visualizing 3D images as axial slices, with the subject's anterior aspect at the top of the screen and the subject's left lateral side on the right side of the screen. I chose to use this convention for visualizing CT data on grab-a-slice, which corresponds to registering the origin of the image in D with the origin of W . Typical data sets are axial slices through CT data, measuring $512 \times 512 \times N$ voxels (where N is the number of axial slices). Using the slice thickness and within-slice inter-voxel distance for a particular 3D image allows for generation of an image-specific transform $\mathbf{T}_{W \rightarrow D}$. This transform ensures that each point in the 3D image is uniquely mapped to a corresponding point in physical space.

I have developed software in the OpenGL environment to render and display images and graphics in 2D and 3D. OpenGL's flexibility permits definition of arbitrary coordinate systems for graphics. To preserve the intuitive nature of the coordinate systems, the graphics coordinate

system, G , is initially registered with the world coordinate frame W . For the psychophysical experiments described below, all rendering is performed orthographically at depth $z = 0$ (the plane of the screen); $\mathbf{T}_{W \rightarrow G}$ is then updated to permit correct rendering of graphical overlays as the grab-a-slice display is moved through the dataset. Later, as described below, I switched to perspective rendering for the stereo display.

Finally, the last coordinate frame is T , the texture memory addressing system. $\mathbf{T}_{D \rightarrow T}$ is a simple scaling transform that maps the image data (typically $512 \times 512 \times N$) into a cube of texture memory with floating-point indices ranging from 0 to 1. The texture memory is the digital storage for the voxel information in the graphical processing unit, permitting rapid hardware-based slicing of the data in arbitrary orientations.

3.7 STEREO DISPLAY AND 3D GRAPHICS COORDINATE SYSTEM

As described above, the second generation of grab-a-slice utilizes a stereo-capable display. The display operates at 120Hz with a pair of wireless LCD shutterglasses, which are controlled by an infrared emitter connected to the graphical processing unit (NVIDIA 3D Vision, Santa Clara, CA). The display shows separate left and right images to each eye of the viewer, as the shutterglasses are darkened synchronously. Following the completion of psychophysical experiments, which were restricted to 2D graphics, the next aim of the thesis research was to create a 3D graphical environment for display of augmented medical image data.

In order to create a stable 3D rendering, the position of the viewer must be tracked simultaneously with the position of the grab-a-slice display. Further, the method of projection in

OpenGL must be changed from orthographic to perspective projection, requiring a redefinition of the graphics coordinate system G with its origin at the location of the viewer (Figure 3.6).

The stereo display shows separate left and right images to the viewer, so separate graphics coordinate systems (G_L and G_R) are identically developed in the software, although for the purposes of the description here, we will assume a single viewpoint (mono). The viewer's eye is at the bottom left corner in Figure 3.6.

An array of six IRED markers was rigidly mounted onto the LCD shutterglasses (Figure 3.5). A circuitboard connecting these markers was fashioned such that the Optotrak camera could simultaneously track the position of the glasses and the grab-a-slice display. As before, the NDI commercial software was used to characterize a "rigid body file" describing the physical arrangement of the six IREDs. A one-time calibration using an optically tracked stylus was then used to determine the location of the viewer's eye while wearing the shutterglasses, which is approximately 2.5cm behind the lens. These positions (one for each eye) are used as the new origin for G , and are tracked along with the six physical markers.

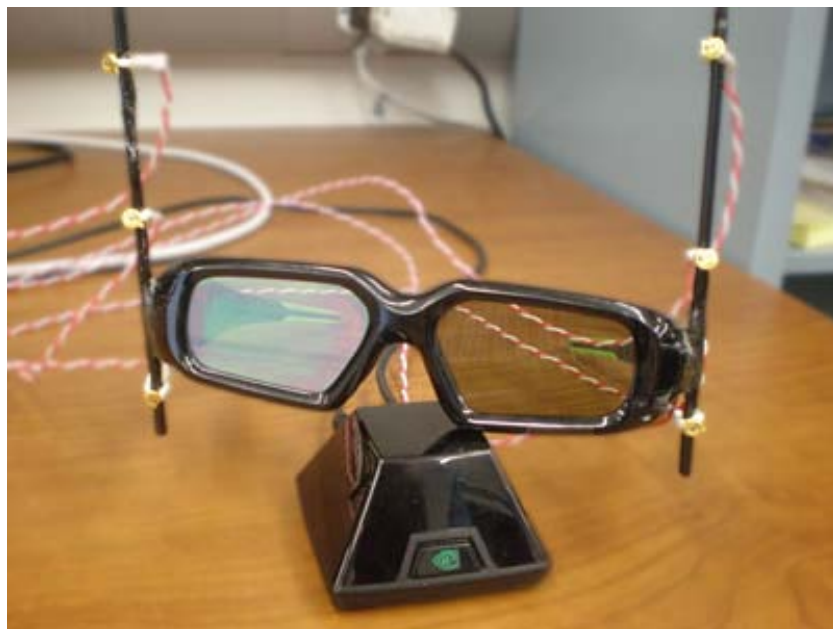


Figure 3.5: 3D shutterglasses with IRED markers rigidly mounted.

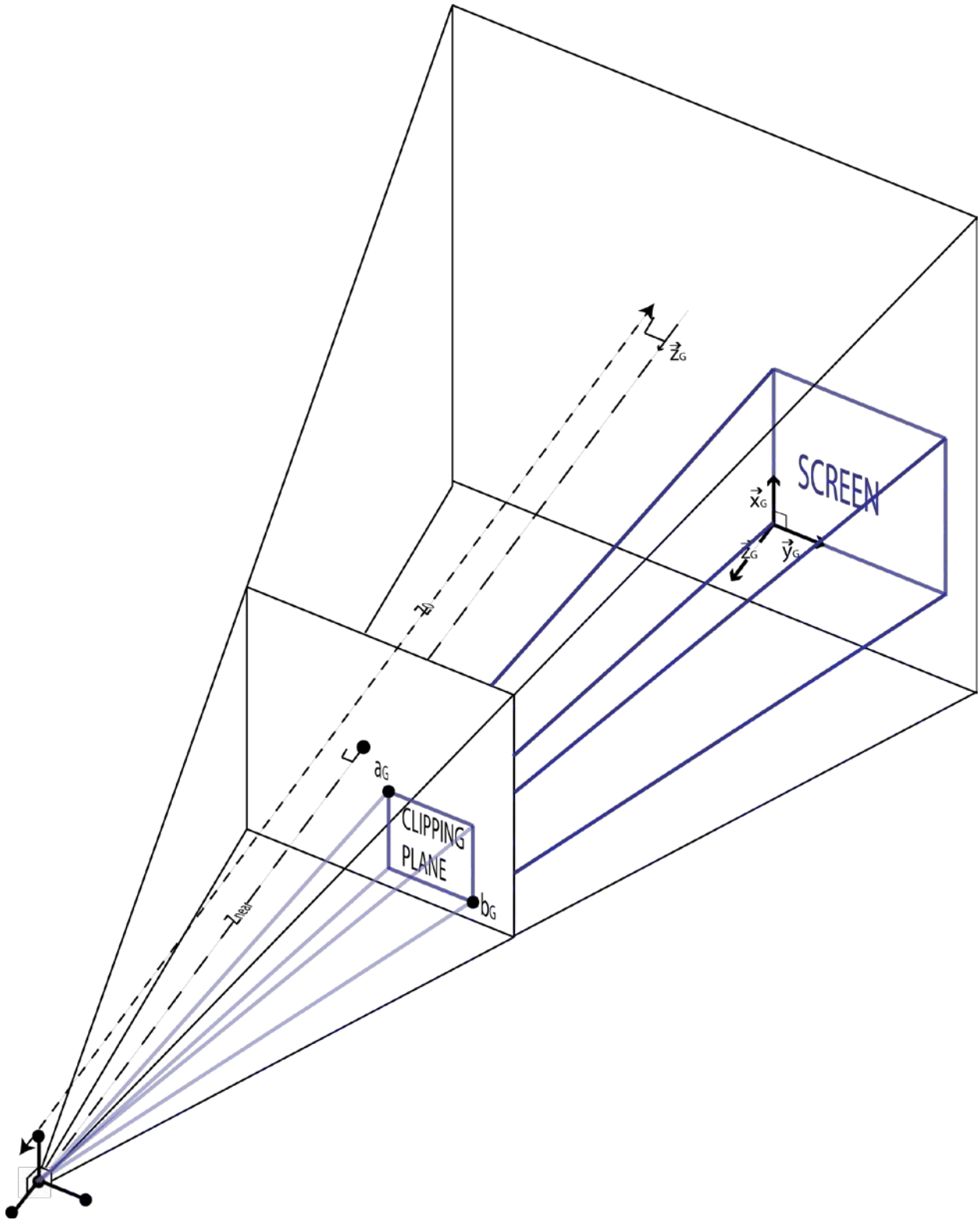


Figure 3.6: GL viewing frustum (blue) for 3D rendering in OpenGL.

Next, we establish the coordinate axes for G . The z -axis is defined to be normal to the plane containing the grab-a-slice display, with the positive direction pointing out of the screen toward the viewer. The x -axis and y -axis are parallel to the line segments demarcating the boundaries of the screen (Figure 3.6).

To generate $\mathbf{T}_{W \rightarrow G}$, I again use four non-coplanar points with known coordinates in G and W . In this instance, I use points near the new origin of G : the new origin and a tripod of points along the new x -, y -, and z -axes. Following the same process described above, a matrix inversion is calculated to yield $\mathbf{T}_{W \rightarrow G}$, and a unique transformation matrix is generated for each eye ($\mathbf{T}_{W \rightarrow GL}$ and $\mathbf{T}_{W \rightarrow GR}$). Any location can then be transformed from its world coordinates into OpenGL graphics coordinates.

I define a viewing frustum (blue in Figure 3.6) of space in which all rendered objects will reside; it is a set of six clipping planes for the rendering. Its location in world coordinates changes with both the location of the viewer and that of the screen; it can be thought of as the viewer's "field of view" of the world as seen through the "window" of the screen. The frustum's range from the viewer is specified with two clipping planes that are normal to the z -axis. Their distances from the viewer (and from the origin in G) are denoted as z_{NEAR} and z_{FAR} . (The far clipping plane is not shown in the figure, as it is chosen to be well beyond the position of the screen and any rendered objects.) Note that in most stereo rendering the z -axis is assumed to pass through the center of the screen, where with the special case of tracking the viewer for an oblique viewing angle, the z -axis may miss the screen entirely. The near and far clipping planes are, nonetheless, chosen to be parallel to the plane of the screen. I choose z_{NEAR} to be a small distance in front of the viewer (30cm). Points \mathbf{a}_G and \mathbf{b}_G (corresponding to the upper-left and lower-right corners of the screen, respectively) specify the lateral boundaries of the near clipping plane, and

therefore the locations of the four lateral clipping planes. The height and width of the near clipping plane are thus determined by proportions; since the image is to fill the entire screen, the clipping plane must have the same proportions as the screen. Because the graphics coordinate space is initialized to have the z -axis normal to the plane of the screen, the x - and y - coordinates of \mathbf{a}_G and \mathbf{b}_G are proportional to the x - and y -coordinates of the screen; they are scaled by the ratio z_{NEAR} / z_0 , where z_0 is the present distance from the origin to the plane of the screen.

The specification of points \mathbf{a}_G and \mathbf{b}_G define the viewing frustum in the OpenGL software. I can then place simple objects into the scene by specifying their locations in W . As described above, the software calculates a new $\mathbf{T}_{W \rightarrow G}$ during each iteration of the code based upon the location of the viewer and the screen, and determines the coordinates in G for each eye. The graphics card then determines how to render the objects to the screen based upon the new coordinates. Generation of binocular disparity is accounted for by the use of separate G coordinate frames for each eye, resulting in a stable 3D rendering of an object outside the plane of the screen.

4.0 PSYCHOPHYSICS EXPERIMENT: EVALUATION OF VISUALIZATION AND NAVIGATION USING GRAB-A-SLICE

As described above, grab-a-slice builds on a series of *in-situ* imaging devices developed in our laboratory that seek to take advantage of the perceptual benefit of placing a tomographic image into a 3D spatial context. With these devices, our lab has explored how novices and experts utilize perceptual cues as well as cognitive processes to perform image-guided tasks. This avenue of research motivated experiments to evaluate how novices interact with 3D medical images using grab-a-slice as compared with a conventional display, both in terms of performing a tracing task on complex vascular structures, as well as estimating the spatial relations between vessel structures in 3D space. We hypothesized that performance using grab-a-slice would be better than that using the conventional display in a variety of measures, as described below.

4.1 METHOD

4.1.1 Subjects

Thirteen naïve observers and three coauthors who were naïve to medical images, totaling four females and twelve males, participated in the experiment. All subjects had normal or corrected-

to-normal vision in both eyes. All gave informed consent as directed by an IRB approved protocol according to which these experiments were performed.

4.1.2 Stimuli

A set of 18 anonymized contrast-enhanced computed tomography images of the thorax was acquired. An expert user identified the pulmonary vasculature in each image and used colored spheres (radius 5mm) to label three structures in each image: the pulmonary artery (PA) just as it exits the right ventricle of the heart (green sphere), the left atrium (LA) where the pulmonary veins drain (blue sphere), and one distal branch of a pulmonary vessel (red sphere) being either an artery or a vein. The spheres were visualized as cross-sectional disks through the sphere overlaid on the particular slice from the data being displayed. Using each image twice, once with the red sphere in an artery and once in a vein, we generated a stimuli set of 36 uniquely labeled vessels. The labeled vessel in each image was connected to either the PA or the LA, but not both. In general, these vessels can be identified in CT images only by their anatomical connection to the heart, rather than pixel intensity or local vessel morphology, so, especially to a novice, they are indistinguishable except by tracing along the vessel, slice by slice, to either the PA or LA.

The distance (in the axial dimension of the image) between the red sphere (marking the unknown vessel) and the correct endpoint (either the PA or LA) was used to classify each vessel into one of three categories, according to its distance from the endpoint: short (5-30 mm), medium (50-80 mm), or long (>100mm). Across images, the distances were uniformly distributed among these categories.

4.1.3 Design

A 3 (Distance) x 2 (Viewing condition: *in-situ* or *ex-situ*) design was implemented. *In-situ* visualization here refers to visualization of the image data on the movable grab-a-slice display, while *ex-situ* refers to visualization on a fixed conventional display. The *ex-situ* display was identical to the *in-situ* display, except that it was stationary on a table directly adjacent to the space in which the grab-a-slice display would be manipulated. Six trials were performed in each of the conditions; a uniquely labeled lung image was used for each of the resulting 36 trials. Trials were blocked into two sets of 18 trials by viewing condition, with the presentation order of trials counterbalanced across blocks and subjects using a Latin-square. The testing order of the two viewing conditions was also counterbalanced to avoid bias from learning, with half of the subjects using *in-situ* first and half using *ex-situ* first.

4.1.4 Procedure

Subjects performed the experiment in a dark room free of peripheral visual cues. Two tasks were performed on each trial.

4.1.4.1 Navigation task

The first task involved navigation through the data set. With the display blanked out, subjects first moved the grab-a-slice display until they found the red sphere in 3D space, which appeared in cross-section on the screen when the display was moved through it. After subjects found the red sphere, the corresponding slice of CT data appeared on the screen, with the red sphere depicted inside an unknown pulmonary vessel. The subjects then followed the vessel by

moving the grab-a-slice display while maintaining a continuous path from the starting point to the endpoint. Eventually, the vessel terminated at the PA (marked with a green sphere) or LA (marked with a blue sphere), at which point the subject made a forced-choice selection about which endpoint was connected to the starting point, by means of a color-coded keypad. Subjects were timed during this “tracing” task, and their accuracy was recorded.

4.1.4.2 Visualization task

The second task, visualization of spatial relations, was performed during and after the navigation task. Subjects were instructed to remember the 3D location of the starting red sphere in space. Upon finding and selecting the endpoint sphere (blue or green), the screen was again blanked out, and subjects were asked to indicate in 3D the vertical plane containing the centers of both spheres. To demonstrate the perceived plane, subjects rotated the blank grab-a-slice display about the y-axis (see Figure 3.3) such that it was deemed parallel to the vertical plane connecting the starting and endpoint locations. Valid angles ranged from -80° to $+80^{\circ}$ relative to the starting position.

4.1.5 Training

Prior to the experimental trials, subjects performed two to four sample trials in each viewing condition to familiarize themselves with the use of the grab-a-slice display and the method of performing the experimental trials. The images used in the sample trials were different from the images used in the experimental trials in that the targets (red, blue, and green) were placed in different locations from any in the test set. Verbal coaching of the subject was helpful

to ensure that each subject understood the two tasks sufficiently to perform the experiment reliably, although quantitative feedback was not provided.

4.1.6 Data Analysis

Paired *t*-tests and ANOVAs were performed to study the effects of viewing condition, order of stimuli, and target distance on accuracy and time-to-completion of the tracing task as well as perceived angle in the visualization task.

Trials in which subjects made errors in choosing the correct vessel during the tracing task were excluded from the analysis of the visualization task. Also, trials in which the subject made a reversal error in the visualization (correct in approximate magnitude but reversed in sign, e.g., they reported a +60° angle for a stimulus that was -60°) were excluded from the analysis of the visualization task. However, these errors were recorded and their frequency was analyzed.

For the visualization task, the slope and correlation coefficient of the least-squares regression line relating stimulus angle to response angle was calculated for each subject (see Figure 4.1 for example), providing a quantifiable measure of performance of each subject across all trials. Perfect performance would have all data points falling along the line $y=x$. Absolute errors (and their mean and standard deviation) were also calculated for each trial.

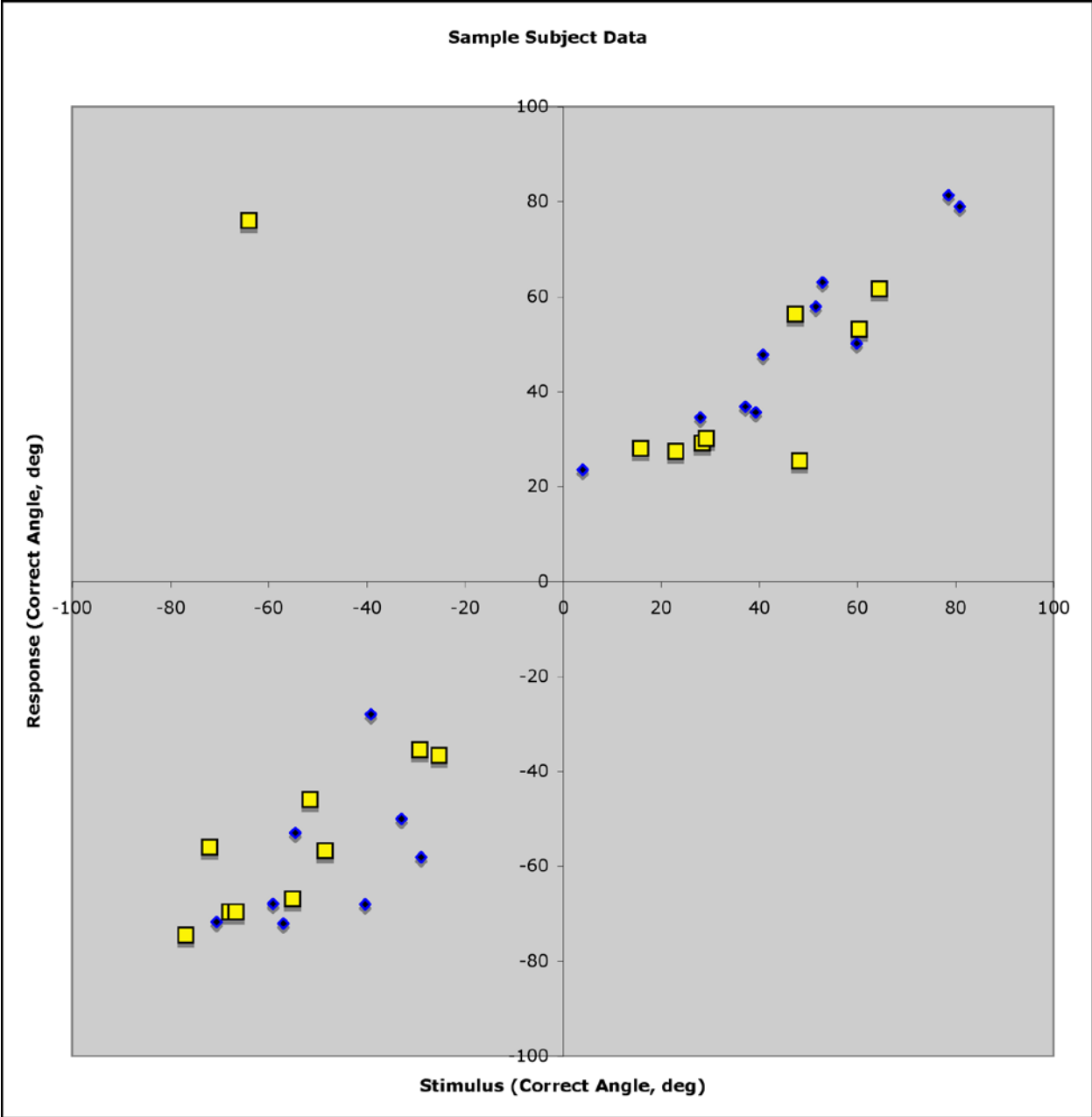


Figure 4.1 Sample subject data for visualization task using grab-a-slice (yellow squares) and conventional display (blue diamonds), showing subject response angles vs. true stimulus angles (as measured in degrees). Note the reversal error in the second quadrant.

4.2 EXPERIMENTAL RESULTS

4.2.1 Navigation task

Subjects were significantly more accurate in the navigation task of identifying the correct endpoint of the unknown vessel when using the grab-a-slice *in-situ* display for visualization ($95.8 \pm 0.3\%$) as compared with using the conventional *ex-situ* display ($93.1 \pm 0.5\%$) ($p < 0.03$). When reversals are included, subjects made fewer errors using grab-a-slice ($10.4 \pm 6.7\%$) as compared with using the conventional display ($13.5 \pm 9.1\%$), though not significantly so. The difference in accuracy was observed without a significant difference in time-to-completion (34 ± 21 s for *in-situ*, 38 ± 14 for *ex-situ*), suggesting that no additional time was needed to achieve the increased accuracy.

4.2.2 Visualization task

Mean absolute errors in the reported angle between starting locations and endpoints were significantly reduced in the grab-a-slice *in-situ* condition ($10.8 \pm 3.8^\circ$) as compared with the conventional *ex-situ* condition ($13.0 \pm 4.7^\circ$) ($p < 0.02$).

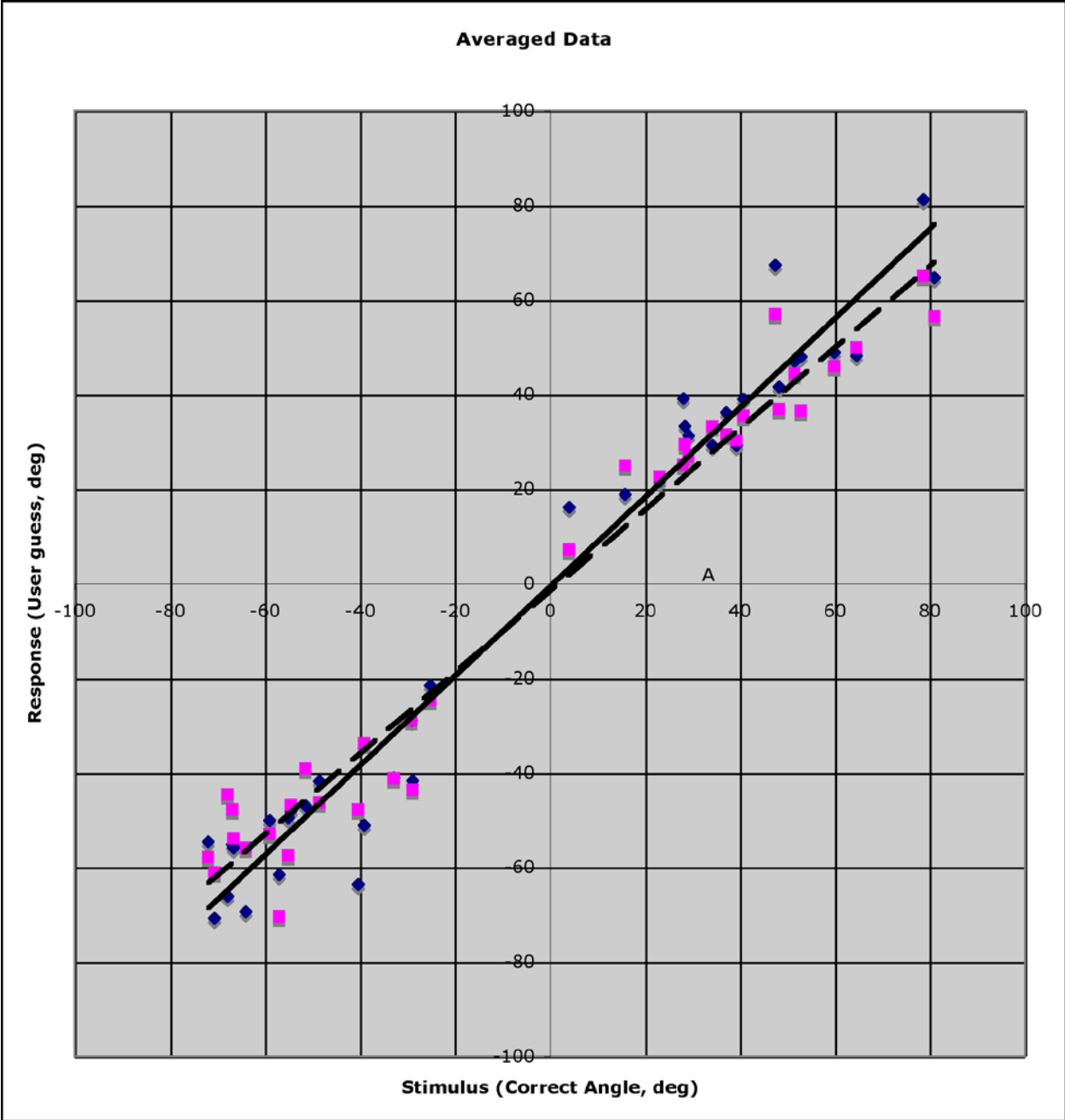


Figure 4.2 Average response for each stimulus angle across all subjects, along with best-fit lines. Solid line: grab-a-slice display $y = 0.95x - 0.58$. Dashed line: conventional display, $y = 0.86x - 1.51$.

The slope of the least-squares regression line relating reported angle to stimulus angle was also compared across displays. The ideal slope for each subject should be 1.0. When the regression was fit to the average response to each stimulus, the slope for grab-a-slice (.946) was closer to 1.0 than the slope for the conventional display (.859), as shown in Figure 4.2. A paired *t*-test between conditions, using slopes calculated for individual subjects, showed the difference was significant ($p < 0.02$). Individual subject results for the pointing task are shown in Appendix A (page **Error! Bookmark not defined.**).

Paired *t*-tests were also performed to assess differences in the variances of the absolute errors in perceived target angles to ascertain if one display modality provided more consistency than another. While absolute error magnitude was significantly reduced with grab-a-slice, the variability in those errors was not significantly different. Further, the correlation coefficients of the least-squares regression lines were also not significantly different between grab-a-slice and the conventional display.

ANOVA was applied to the data to look for effects of target distance, stimulus order, and modality order. No significant effects were observed.

4.3 SUMMARY AND CONCLUSIONS OF EXPERIMENT

These data support the hypothesis that the *in-situ* display of medical image data confers a benefit to naïve subjects in both navigation through and visualization of 3D medical image data. Accuracy in the tracing task was significantly higher, absolute errors in the visualization task were lower, and reported spatial relations between vessel structures showed a closer linear

relationship to true values with the grab-a-slice *in-situ* display as compared with the conventional *ex-situ* display.

One interesting finding was that no effect was observed due to the distance from starting point to endpoint; there was no significant difference in performance in either task as a function of this distance. With the number of trials and subjects, we expected that the study was adequately powered to detect significance at the 0.05 level in a generalized linear model. The range of distances in the experiment represented typical clinical distances over which radiologists might be expected to navigate during PE detection; perhaps this range is simply too narrow to further divide into “easier” and “harder.”

Given that the subjects were naïve to studying medical image data, it was a challenge to identify a task that was simple enough for novices to learn, yet difficult enough to adequately measure differences in performance across the experimental conditions. Clearly, the experiments we settled upon were effective in this regard, as they allowed us to demonstrate underlying perceptual differences involved between *in-situ* and *ex-situ* image displays. We have submitted these results for publication in the IEEE Transactions on Visualization and Computer Graphics [57]. Additional tasks, we believe, would expand the psychophysical domain in which *in-situ* display is demonstrably superior.

5.0 STEREO VISUALIZATION RESULTS

The psychophysics experiments described in the previous chapter did not utilize the stereo capabilities of the grab-a-slice display, as they were assessing the movable tomographic display, which did not rely on 3D graphics. Following the completion of those experiments, the final phase of my thesis research focused on the development of a 3D graphical environment within grab-a-slice in which augmentation of tomographic images with out-of-plane structures is possible. The results of this research are described here.

5.1 PRESERVATION OF PREVIOUS FUNCTION OF GRAB-A-SLICE

As described in Chapter 3.7, I changed the rendering system in OpenGL from an orthographic rendering of a 2D slice to a perspective rendering of a 3D scene, incorporating the position of the viewer when calculating the required projection transforms. The first test to demonstrate the success of the perspective rendering was to display the correct tomographic image on the screen just like any other 3D object in the scene. Essentially, I needed to duplicate the mono behavior of grab-a-slice in the stereo environment by providing the appropriate different projection of the tomographic image to the left and right eyes.

Recall that one of the most powerful cues for generating the perception of depth is stereopsis, in which the brain receives disparate images from each retina and uses the differences to infer depth. In the area of computer graphics, as noted previously, we can use LCD shutterglasses to facilitate this process with the stereo display by alternately opening and closing each lens synchronously with the display, which shows alternating images at 120Hz (60Hz refresh rate for each image). Thus, the right eye always sees one image, while the left eye always sees another. To prove that the stereo function was working correctly, I started with the mono version of grab-a-slice and introduced a fixed horizontal offset between the left-eye and right-eye images, which had the effect of making the tomographic slice appear to be behind the plane of the screen when the shutterglasses were worn. This demonstrated the effectiveness of out-of-plane rendering with the shutterglasses, but was discarded once I developed the general transform matrices described in Chapter 3.7.

To test the validity of these graphics transforms, the tomographic slice was treated as a graphical object and fed through the transforms to be rendered on the screen at its present location. The fact that there was no disparity between the resulting left-eye and right-eye images demonstrated that the new stereo transforms were working correctly, at least for information in the plane of the screen.

In mathematical terms, the tomographic slice was calculated the same way as previously: I used the world coordinates of the screen to map into the texture memory containing the image, and selected the corresponding slice. The world coordinates of the slice were then put through separate transform matrices $\mathbf{T}_{W \rightarrow GL}$ and $\mathbf{T}_{W \rightarrow GR}$ to yield the graphics coordinates for its location with respect to each eye. I confirmed that these transforms were operating correctly, as the tomographic slice (which resides in the plane of the screen) was identical for both the left-eye

and right-eye images despite their positions being calculated independently (Figure). Thus, the mono version of grab-a-slice is preserved correctly in the stereo environment. The image remains stable as the viewer (wearing the tracked shutterglasses) moves around the scene, provided that the line of sight between the Optotrak camera and the IRED markers on the glasses is not obscured.

5.2 RENDERING OF 3D OBJECTS

The major benefit of using the stereo display is to augment the tomographic slice with 3D information that resides outside of the slice plane. Ergo, the next goal was to create a rendering of a 3D object in space that remains stable as the viewer moves around it. I chose a small black square floating two inches in front of the screen (and in front of the tomographic slice plane) as an initial test object. This was successful – the square has a fixed location in space independent of the viewer’s location and the screen’s position and orientation.

I chose a 3D sphere floating in front of the screen as the next test object. I utilized the OpenGL lighting and shading models to illuminate the sphere, as these models serve as cues that further enhance the perception of a 3D scene. The sphere does remain stable as the grab-a-slice display is moved along the z_G -axis or when the viewer moves horizontally. However, as the angle between the viewer and the display becomes more acute (horizontally and vertically), some distortions become apparent for reasons that are still unsolved. This may be due to the manner in which the triangle mesh of the sphere’s surface is calculated by the graphics library, and the order in which the vertices of these triangles enter the pipeline of graphical transformations. Further, it may be caused by the fact that the precise location of the viewer’s retina is impossible

to incorporate into the scene; we have approximated this location as best we can, but the error in this approximation may be partially responsible for these distortions. Work is ongoing to solve this problem.

Once a sphere is stably rendered sans distortion, the next goal is to put multiple spheres into the scene along with the tomographic slice. Here, we will utilize the OpenGL models of opacity and depth buffering to make 3D objects in front of the screen appear translucent so that the tomographic image appears clearly behind them, as well as to make the tomographic images translucent to objects rendered behind it.

As described in Chapter 2.4, the current version of Shells and Spheres [43] is a fully automated n -dimensional method of finding medial spheres whose union represents a segmentation in a medical image. We have taken the medial spheres generated by this framework and created perspective renderings in an OpenGL window outside of the grab-a-slice environment (Figure). Clearly, these data are ideal for 3D rendering with grab-a-slice. A new graduate student in our lab, Vikas Shivaprabhu, has taken over this part of the research, and plans to use grab-a-slice to render sets of spheres that represent the segmentation of objects of interest. While this will begin with visualization demonstrations, we envision subsequently using grab-a-slice to interact with the segmentation algorithm itself in a supervisory role.

6.0 DISCUSSION AND FUTURE DIRECTIONS

Grab-a-slice in its current form has much promise as a clinical tool and a device for further psychophysical research. Currently it can be used for visualization of the results of 3D medical image analysis routines, and it may serve as a supervision tool for such routines in the future. In the course of the development of the device, we have discovered areas for design improvement and established some of the strengths and weaknesses of grab-a-slice relative to the current clinical methods of visualization, which are discussed below.

6.1 ADVANTAGES AND LIMITATIONS OF GRAB-A-SLICE

The first (tabletop) generation of grab-a-slice used a wooden frame with handles. As mentioned earlier, this frame was bulky and unwieldy, making it difficult to manipulate at times and introducing visual distractors for the user. The second-generation device containing the boom-mount on the back of the display solved many of these limitations. In its current form, grab-a-slice still requires a significant amount of space for operation. The boom arm is mounted on a large heavy cart. An alternative approach could be to mount it to a wall or ceiling. The presence of the Optotrak Certus system for tracking position is not only space-consuming but prohibitively costly for most clinical applications. Some combination of cheaper technologies, including other optical, RF, or inertial tracking systems, may make more sense for clinical use.

The boom itself presents both advantages and limitations. Although it does allow for the operator to let go and leave it in its current location, the non-negligible weight of the boom and angle-dependent effect of joint friction introduced irregular forces that vary somewhat arbitrarily with position of display. This noticeable issue could be resolved by using a tracked, lightweight, handheld tablet device that could wirelessly display the real-time data from a separate computer. The 2011 Consumer Electronics Show included the debut of the G-Slate, a tablet made by LG Electronics (Seoul, South Korea) that could serve as the basis for such a display; its internal gyroscopes could be paired with the rear-mounted stereo cameras to replace the expensive optical tracking system, and its wireless connection could be used to stream video from a nearby computer while that computer's GPU performs the texture mapping. The G-Slate also has stereo display capability for use with shutterglasses similar to those in the current system; the front-facing camera could be utilized to monitor the position of the viewer.

Grab-a-slice evolved from the clinical and virtual sonic flashlight projects, which had the interesting property of creating a virtual image seen through a semitransparent mirror. The presence of the mirror is crucial to the function of those devices, but one drawback is that the image suffers from subtle degradation; when viewed through the mirror, the effective contrast and brightness of the reflected image is decreased. Grab-a-slice doesn't have this problem because the viewer's gaze is focused the display's real image when only a tomographic slice is being displayed. When using grab-a-slice to show stereo images out of the plane of the display, however, the effect is that the viewer is looking through a "window." The tomographic slice resides within the windowpane itself, while other objects appear in front of or behind that windowpane. Thus some form of compromise similar to that of the SF's half-silvered mirror is inevitable in combining the in-plane slice with the out-of-plane stereo graphics. But with

intelligent use of opacity, lighting, and other graphical cues, we hope that significant information may not be lost, leaving the viewer with the perception that the slice plane is moving through data anchored in a 3D location while out-of-plane objects are similarly anchored in the world. In such a system, as with the SF, the perception of 3D shape and location should be facilitated by the fusion of these two sources of information.

6.2 CLINICAL UTILITY

This new method of medical image display shows promise in at least three areas of clinical application.

One area is preoperative planning. Surgeons often examine medical images before performing a procedure. For example, a surgeon attempting to extract a bullet from a patient's abdomen may look at an abdominal CT scan to identify the bullet's location, possible obstacles, and to determine an optimal path of approach. At present, the surgeon must cognitively relate this information, displayed as 2D slices on a stationary screen, to the 3D location of the patient in the operating room. Being able to examine the CT data with grab-a-slice may provide the surgeon with a more intuitive sense of 3D anatomical relationships, analogous to what is termed in the military as, "situational awareness." The surgeon may use it to preoperatively plan and record a surgical path and demonstrate it to the surgical team. The screen may be located on a table directly adjoining the patient, with the "invisible patient" oriented parallel to the real one, thus making corresponding orientations and distances directly comparable.

A second potential clinical application involves using grab-a-slice as a training tool. Medical students, nurses, residents, and other health professionals often have difficulty learning

to interpret 3D medical images; the orientation of the patient and the relationship between slices is not always readily apparent. Medical students in the anatomy laboratory could dissect a liver in the cadaver while examining a grab-a-slice rendering of an abdominal CT scan of the same liver, using a 3D rendering of the segmented liver to guide their dissections and gaining expertise with medical images at an early stage in their training.

A third clinical realm for grab-a-slice is diagnostic radiology. As described above, one particular application we are studying is the diagnosis of pulmonary embolism (PE), an acute, life-threatening, and treatable condition with over 250,000 incident cases in the USA annually. PE is diagnosed by a combination of clinical symptoms, laboratory results, and medical imaging to determine the presence, location, and size of potential emboli. Radiological evaluation of CT data for evidence of PE involves tracing branches of the pulmonary vessels containing suspected emboli back to the heart, to determine whether the vessel in question is an artery or a vein (pathologic emboli are always in the arteries). As we demonstrated in our psychophysics experiments, novices showed better performance with grab-a-slice at vessel-tracing tasks and visualization of 3D structures. Future experiments with expert radiologists could provide meaningful data into the role of learning and prior top-down knowledge in grab-a-slice as well as some constructive feedback about the design of the next iteration of the device.

The role of stereo vision in the clinical setting remains a topic of debate [58]. Stereo display of medical images has been explored for each of the above applications, as well as a fourth one – minimally invasive surgery – but has yet to be adopted as a mainstream technique for visualization, primarily because the studies have not been extensive or decisive. For example, in preoperative treatment planning, virtual reality systems using perspective renderings of 3D images have been used, and one group has investigated stereoscopic imaging for planning

of radiation therapy [59], but despite the potential benefit of direct perception of volumes, distances, and angles in 3D images prior to intervention, stereoscopic displays have not been pursued further for this application. In diagnostic radiology, studies have attempted to evaluate the potential benefits for angiography [60][61], skeletal imaging [62][63], 3D ultrasound [64], mammography [65], and other diagnostic applications, and while positive trends are seen, significant benefits for stereoscopic imaging have not been consistently demonstrated.

Like many medical specialties, diagnostic radiology is a conservative field, relatively resistant to novel techniques or methods, and with good reason. Generally, new technology is only accepted when it confers a quantifiable benefit to clinical care, in the form of reductions in medical errors, training time, recovery time – and cost. Not only do these benefits have to be demonstrable through rigorous scientific study, they should be immediately apparent to clinicians when they first use the technology – one of the promising characteristics of the sonic flashlight. In light of the effect size seen in the grab-a-slice studies, I hypothesize that the biggest impact for grab-a-slice would be as a training tool. During their training, and before they have had years of experience learning a particular approach, medical professionals may be more receptive to alternative approaches. A clinically viable iteration of grab-a-slice could fill this role in medical school.

6.3 FUTURE PSYCHOPHYSICAL EXPERIMENTS

In addition to offering potentially useful clinical applications, grab-a-slice represents a new platform for at least two areas of fundamental psychophysical research: 3D visualization and 3D navigation. The present work began the evaluation of these areas using in-plane tomographic

images; adding stereo graphics to render 3D objects around the image plane makes more experiments possible.

Grab-a-slice follows a series of *in-situ* display devices developed in our laboratory. We have consistently shown a perceptual and cognitive benefit when using *in-situ* images, and the results presented here suggest that grab-a-slice shares many of these same benefits. The processes that humans use to build up mental representations of 3D structures are of great interest to the psychophysical community. Grab-a-slice could be used to study how humans perceive curvature in 3D and to evaluate the ability to define a path through a 3D maze structure. Navigation through a maze is analogous to a surgeon's preoperative and intra-operative route planning, avoiding obstacles to reach targets. Thus the *in-situ* nature of grab-a-slice provides a novel way in which to study these human perceptual and cognitive processes, and compare them to those used with conventional displays.

6.4 COMPUTER-AIDED SEGMENTATION ALGORITHMS

In our ongoing development of grab-a-slice, we are developing methods of graphical augmentation using stereo display to improve the ability of users to understand the raw content of a tomographic slice in the context of the surrounding 3D anatomy and to improve their ability to navigate through a 3D dataset. The current state of the device allows for a 3D sphere to be rendered (albeit with some distortion) at a position between the display and the viewer, with programmable illumination of the sphere.

Future work will explore the application of grab-a-slice for guidance and supervision of segmentation using a Shells and Spheres algorithm designed to benefit from the human

interaction. We propose that this paradigm may provide a symbiotic coupling of top-down processes (human-user guidance) and bottom-up processes (pixel-level computational analysis) for semi-automated image segmentation.

APPENDIX A

INDIVIDUAL SUBJECT RESPONSE DATA FOR VISUALIZATION TASK OF PERCEPTION EXPERIMENT

Sixteen subjects were used in the perception experiment described in Chapter 4.0 . Their individual responses for the visualization task, described in Chapter 4.1.4.2, are shown here. Response angles are plotted against stimulus angles for both the grab-a-slice (GAS) and conventional (CUS) display conditions. (The acronym “CUS” stands for “conventional ultrasound,” which was the control condition in previous psychophysical experiments on the sonic flashlight; the acronym remained in place for the grab-a-slice experiments even though there is no ultrasound data being used.) If a subject had perfect performance, the data points would lie along the line $y = x$; reversal errors (as described above) are shown in quadrants II and IV.

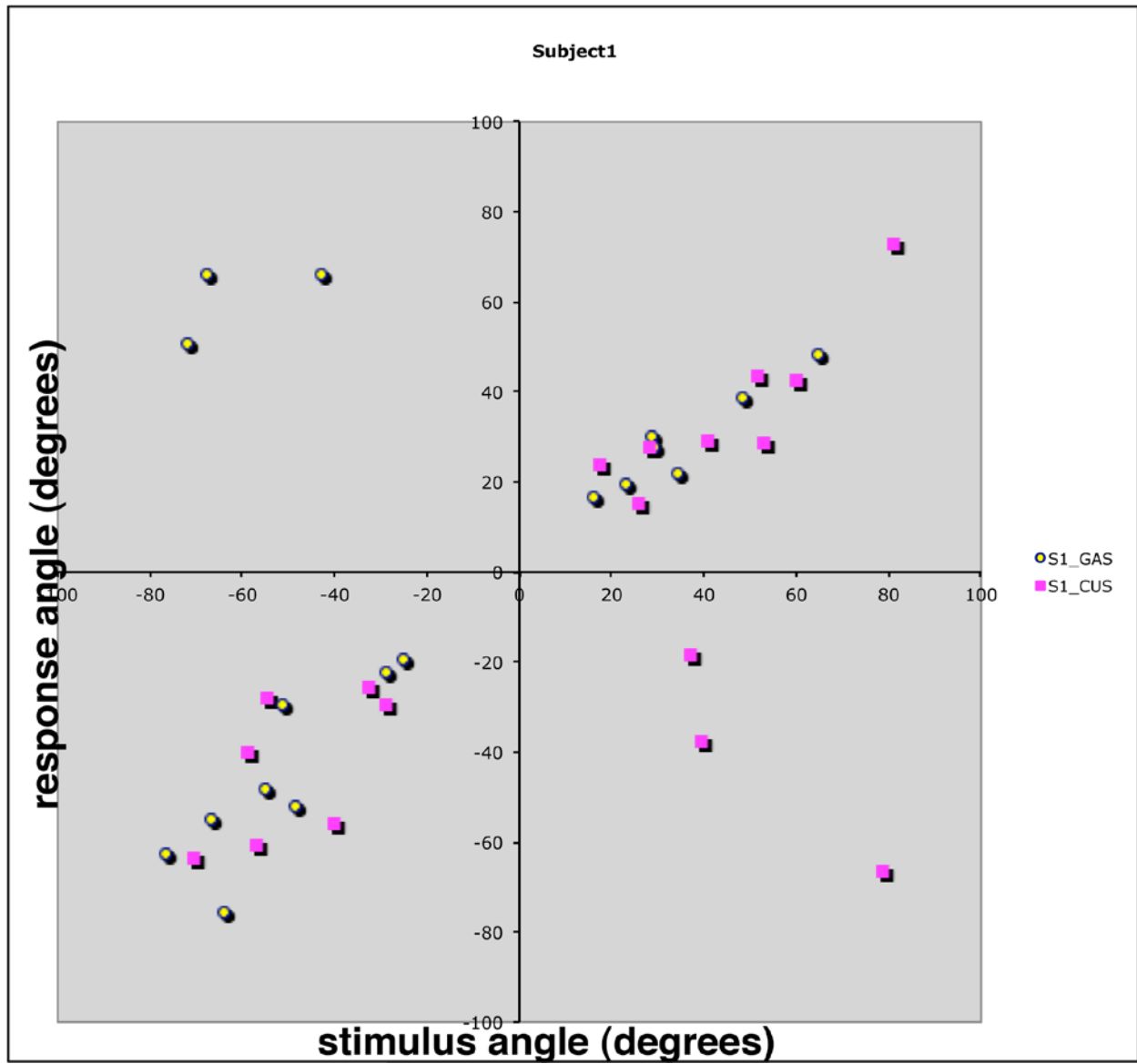


Figure A.1: Visualization data for subject 1.

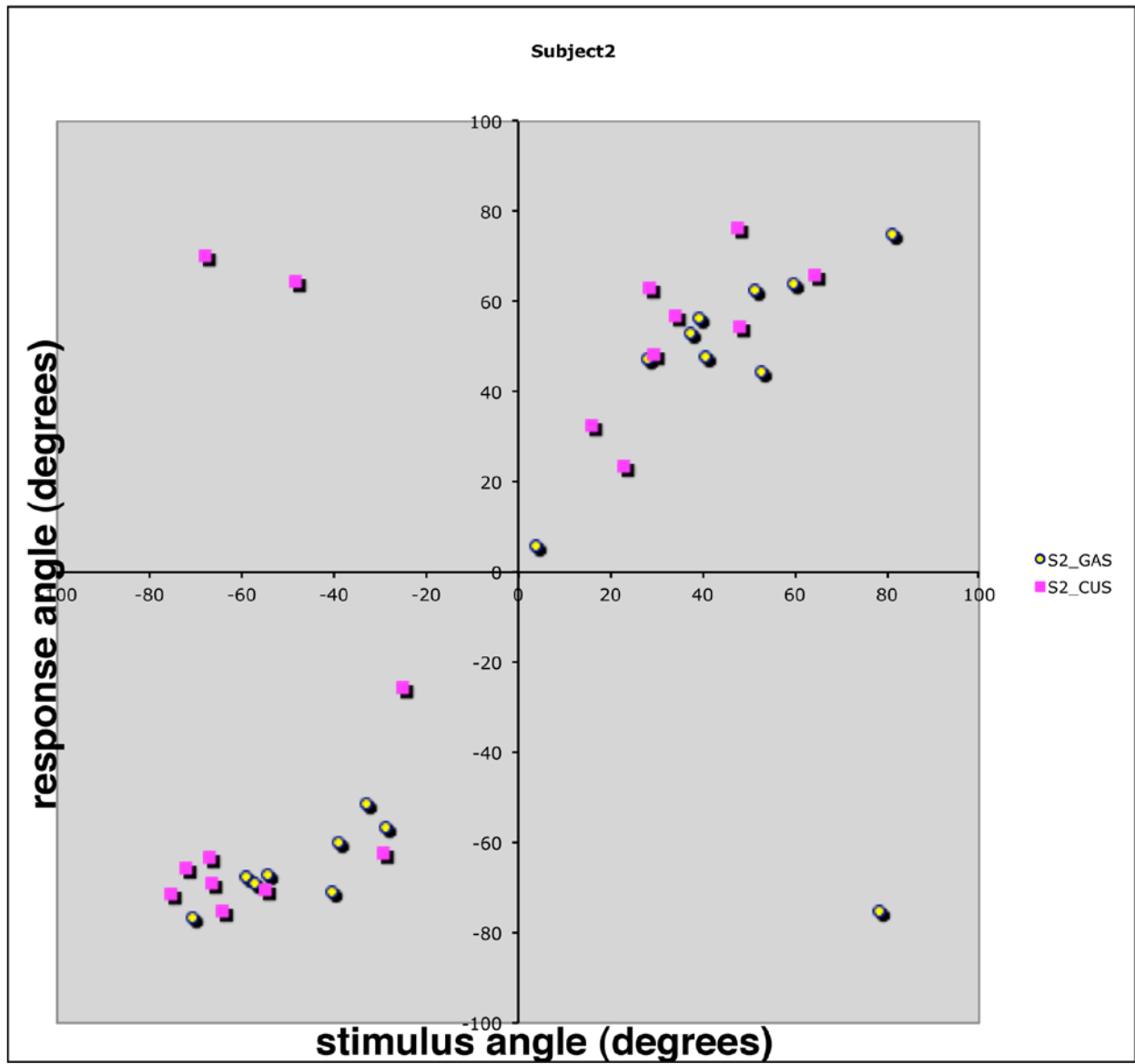


Figure A.2: Visualization data for subject 2.

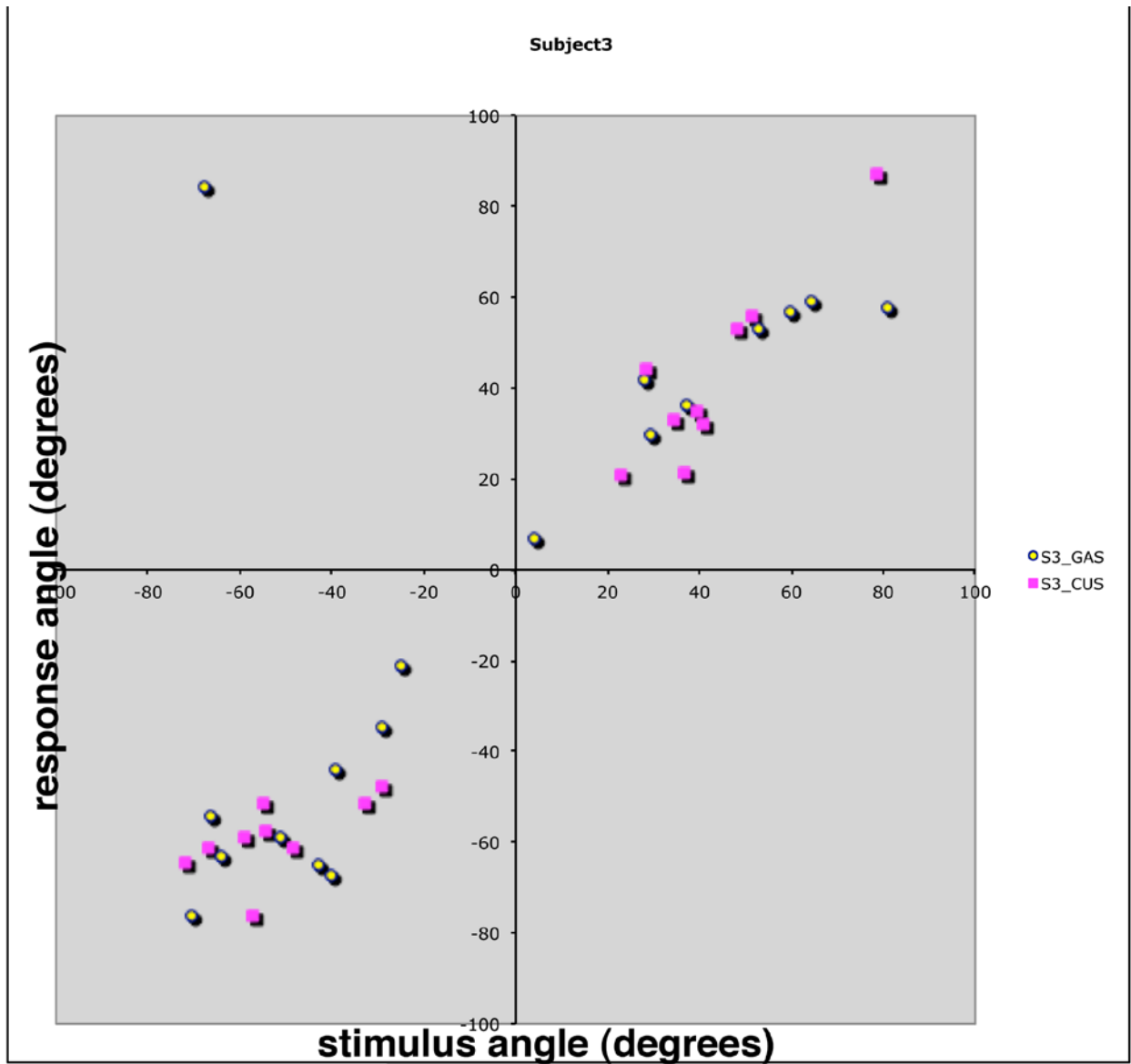


Figure A.3: Visualization data for subject 3.

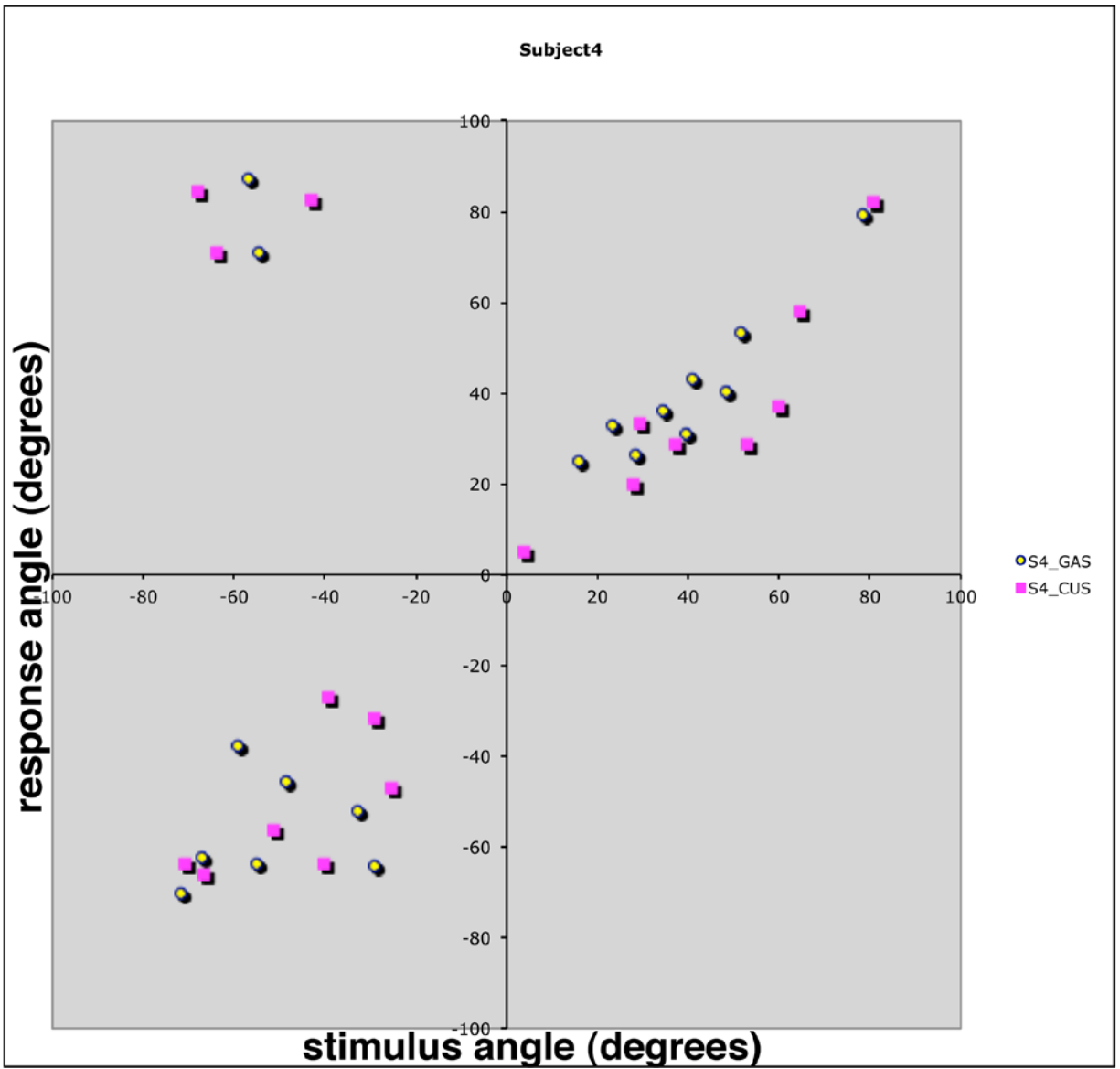


Figure A.4: Visualization data for subject 4.

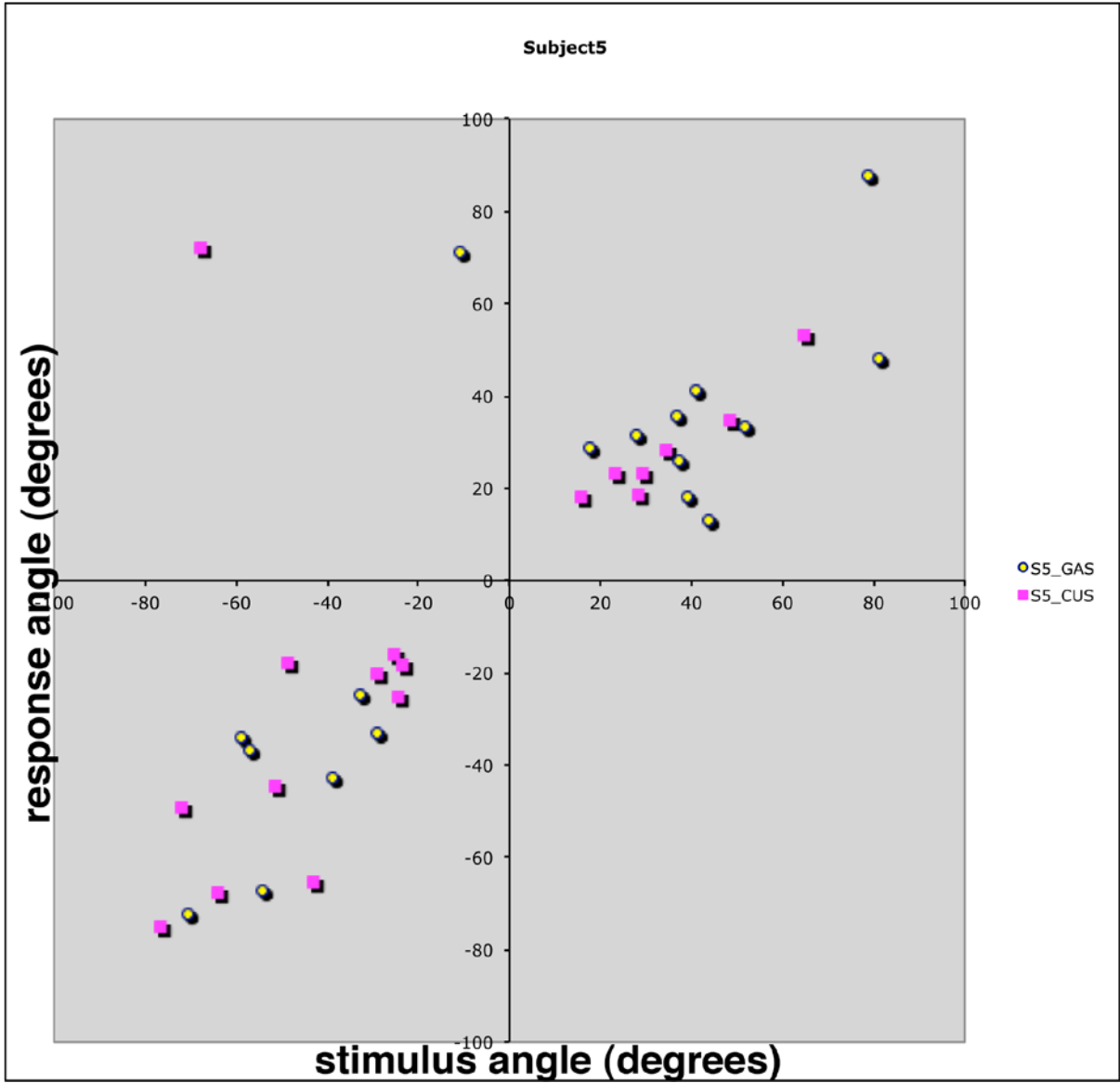


Figure A.5: Visualization data for subject 5.

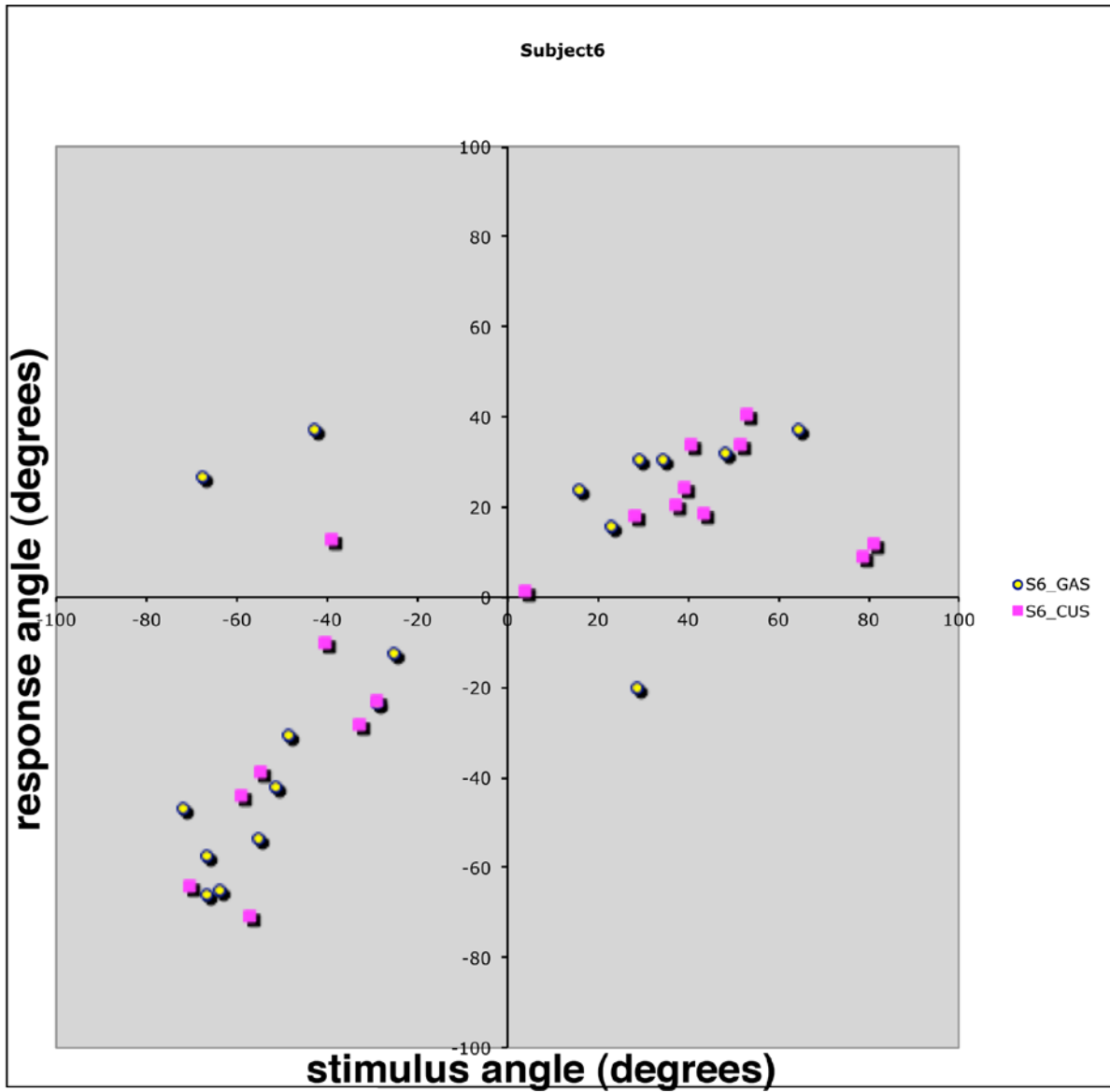


Figure A.6: Visualization data for subject 6.

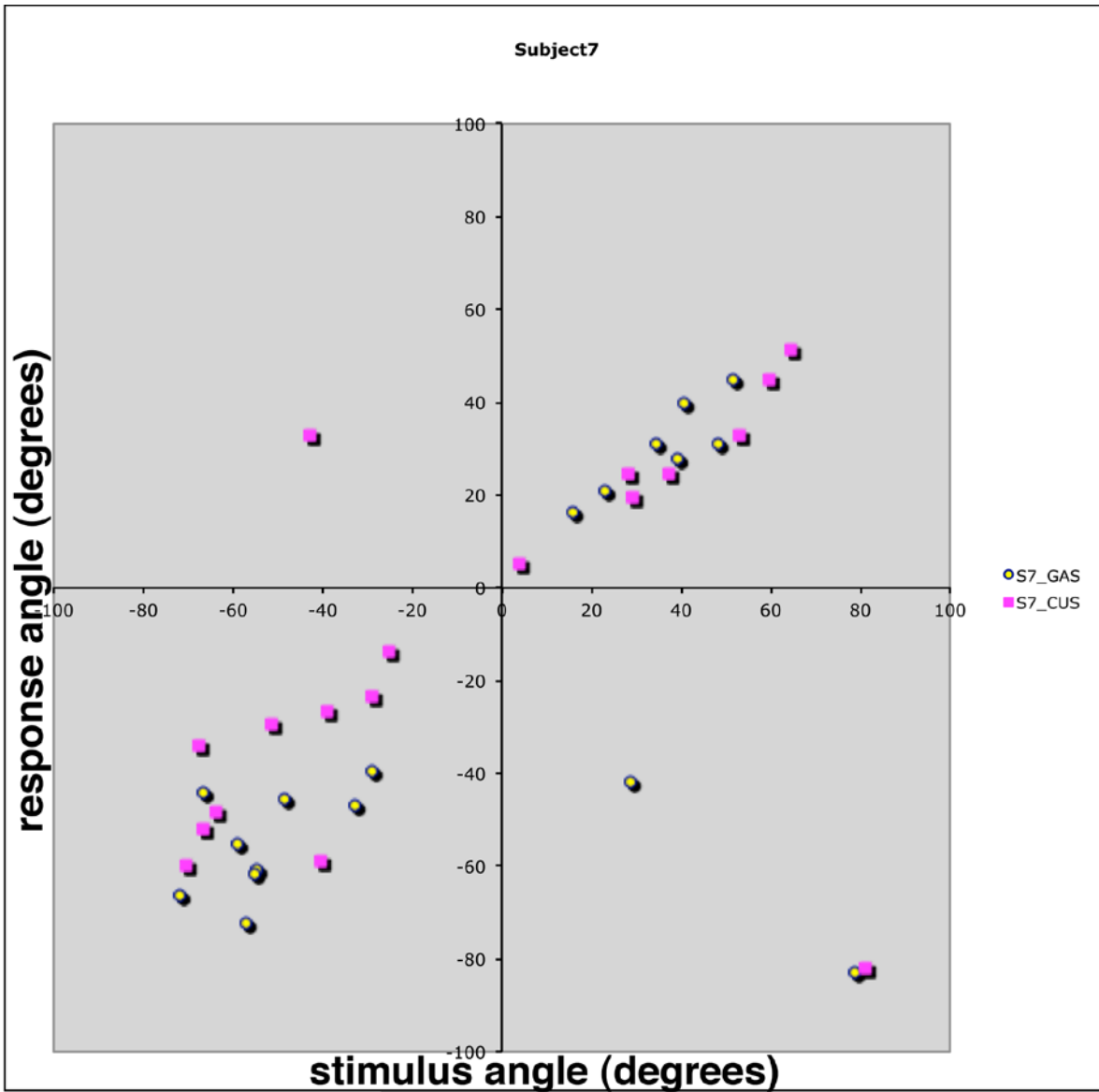


Figure A.7: Visualization data for subject 7.

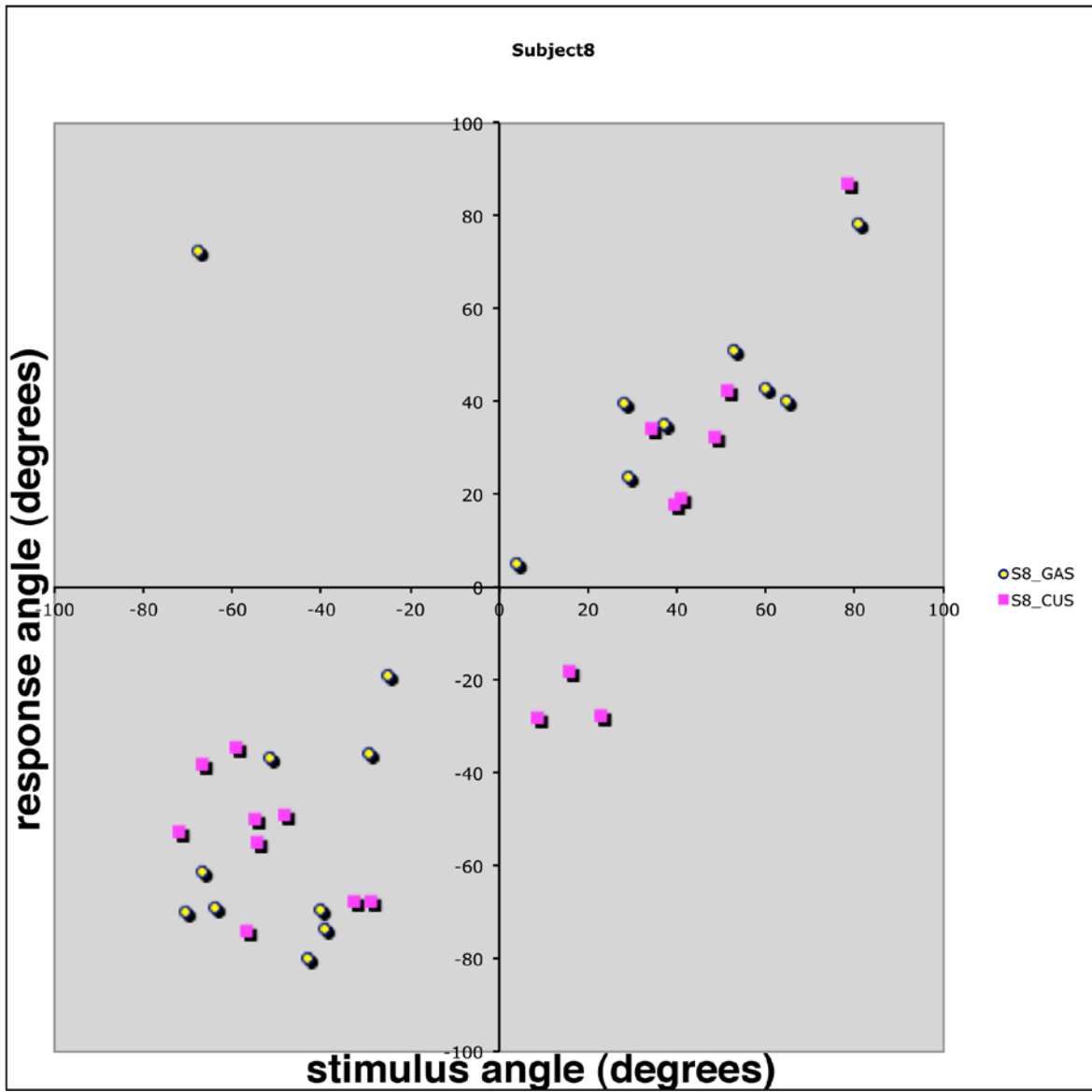


Figure A.8: Visualization data for subject 8.

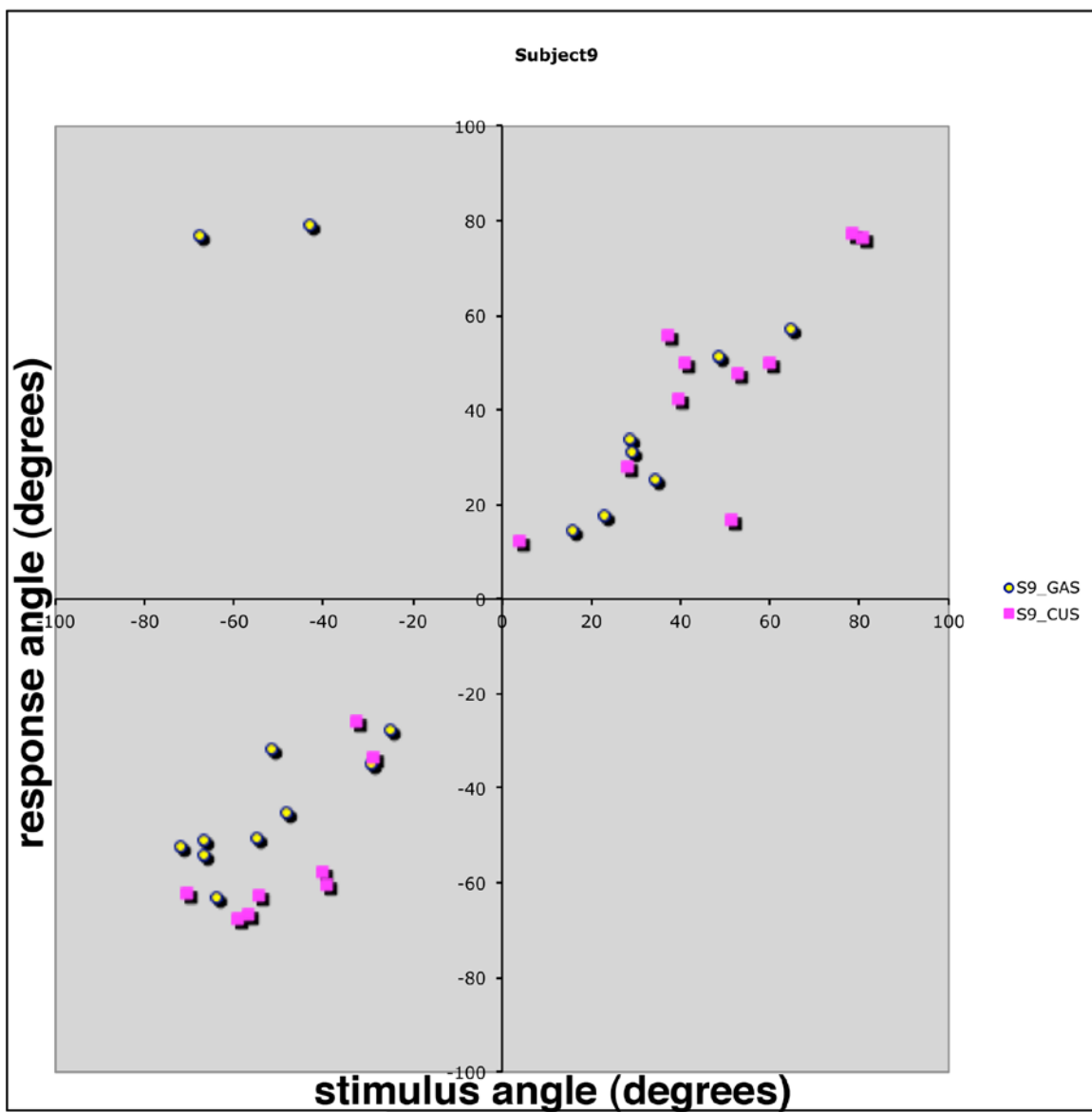


Figure A.9: Visualization data for subject 9.

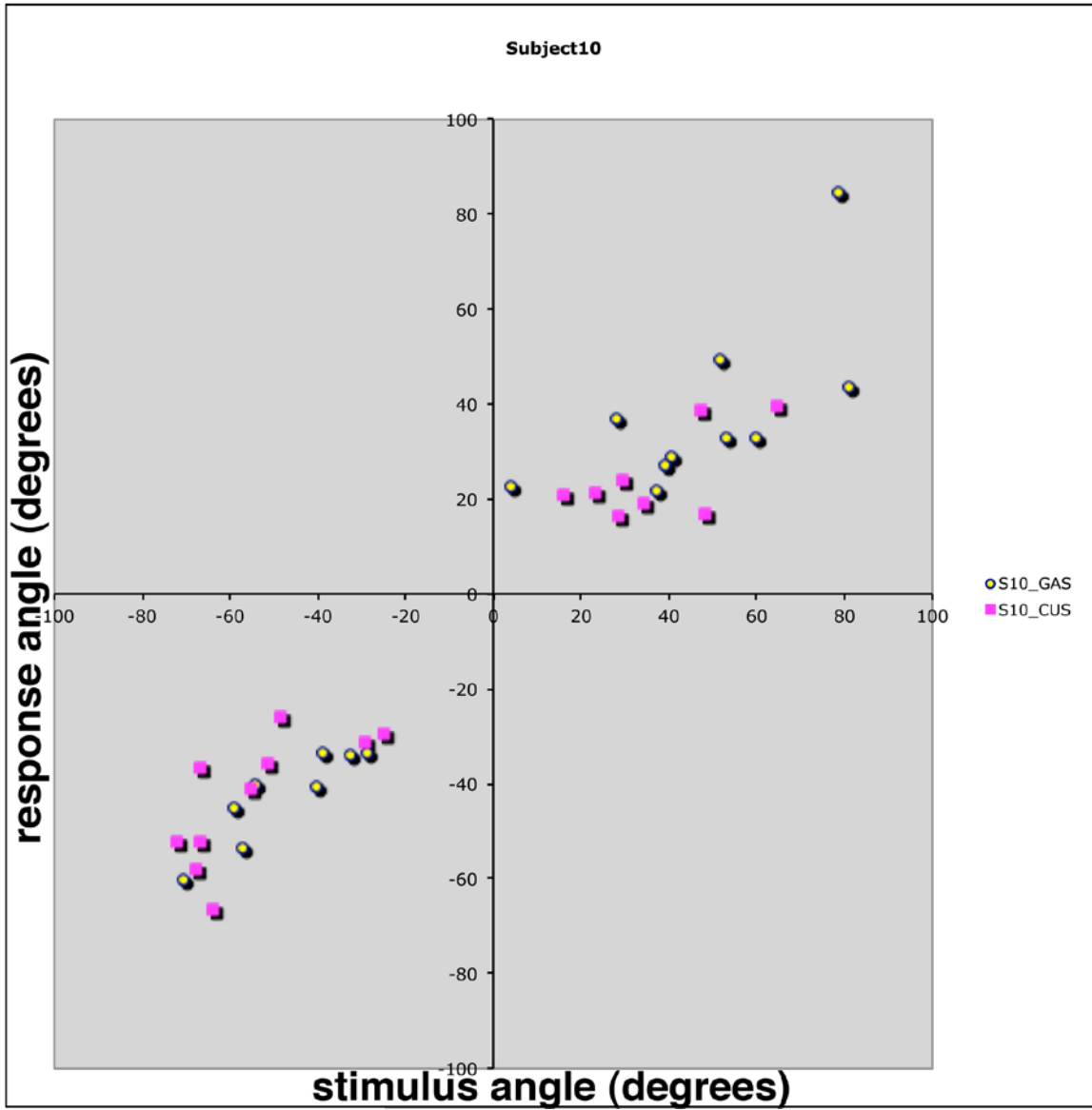


Figure A.10: Visualization data for subject 10.

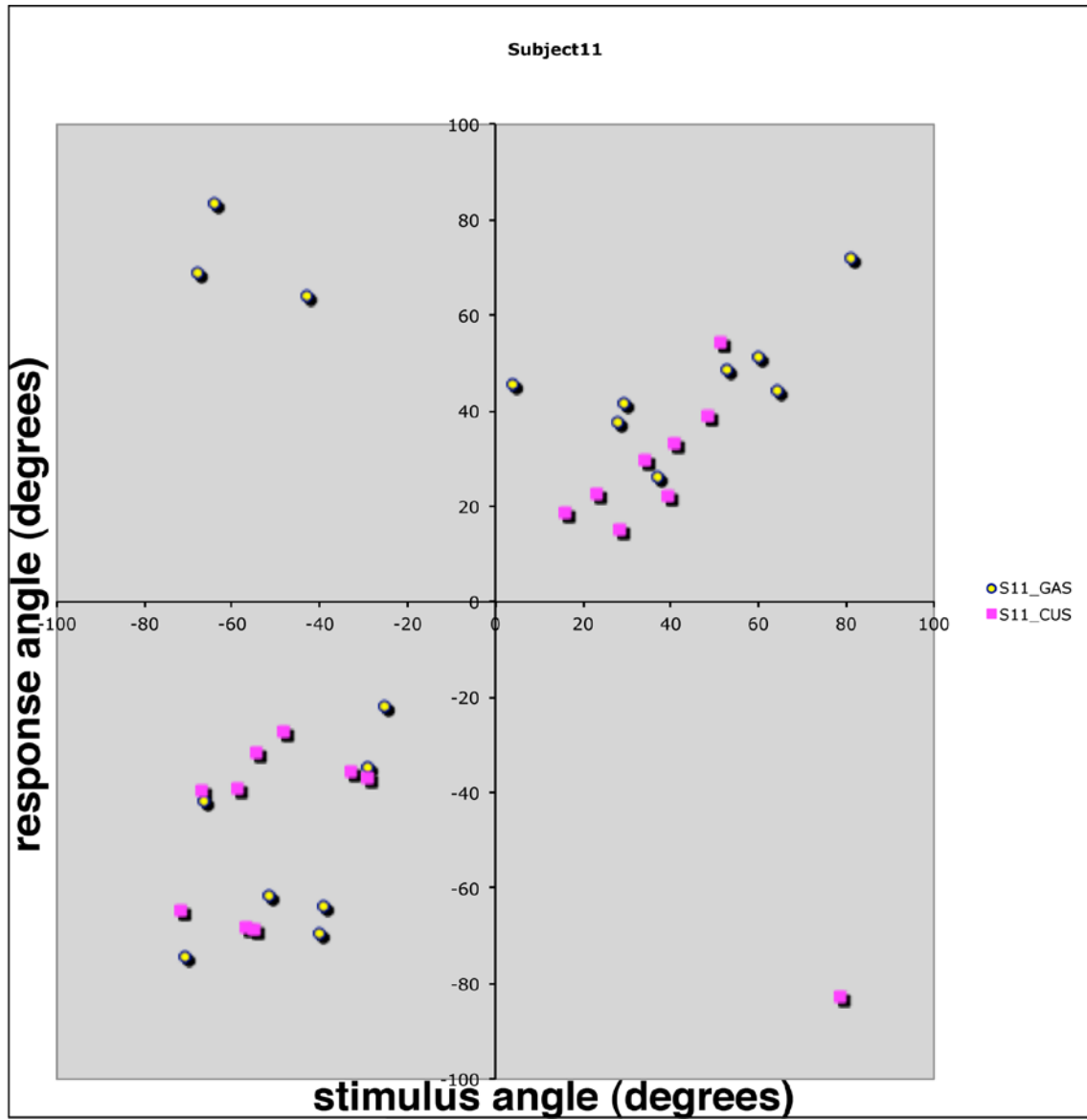


Figure A.11: Visualization data for subject 11.

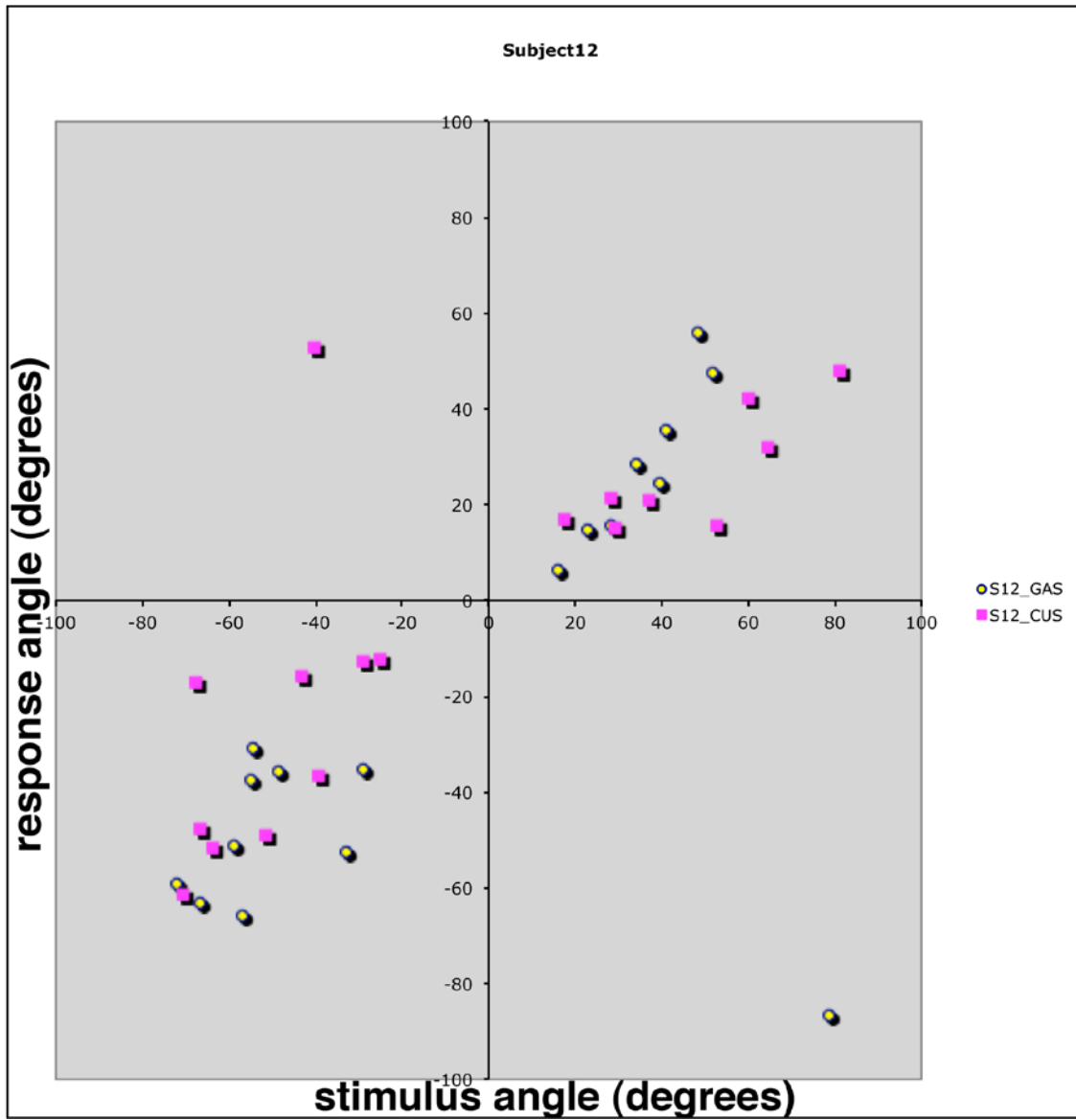


Figure A.12: Visualization data for subject 12.

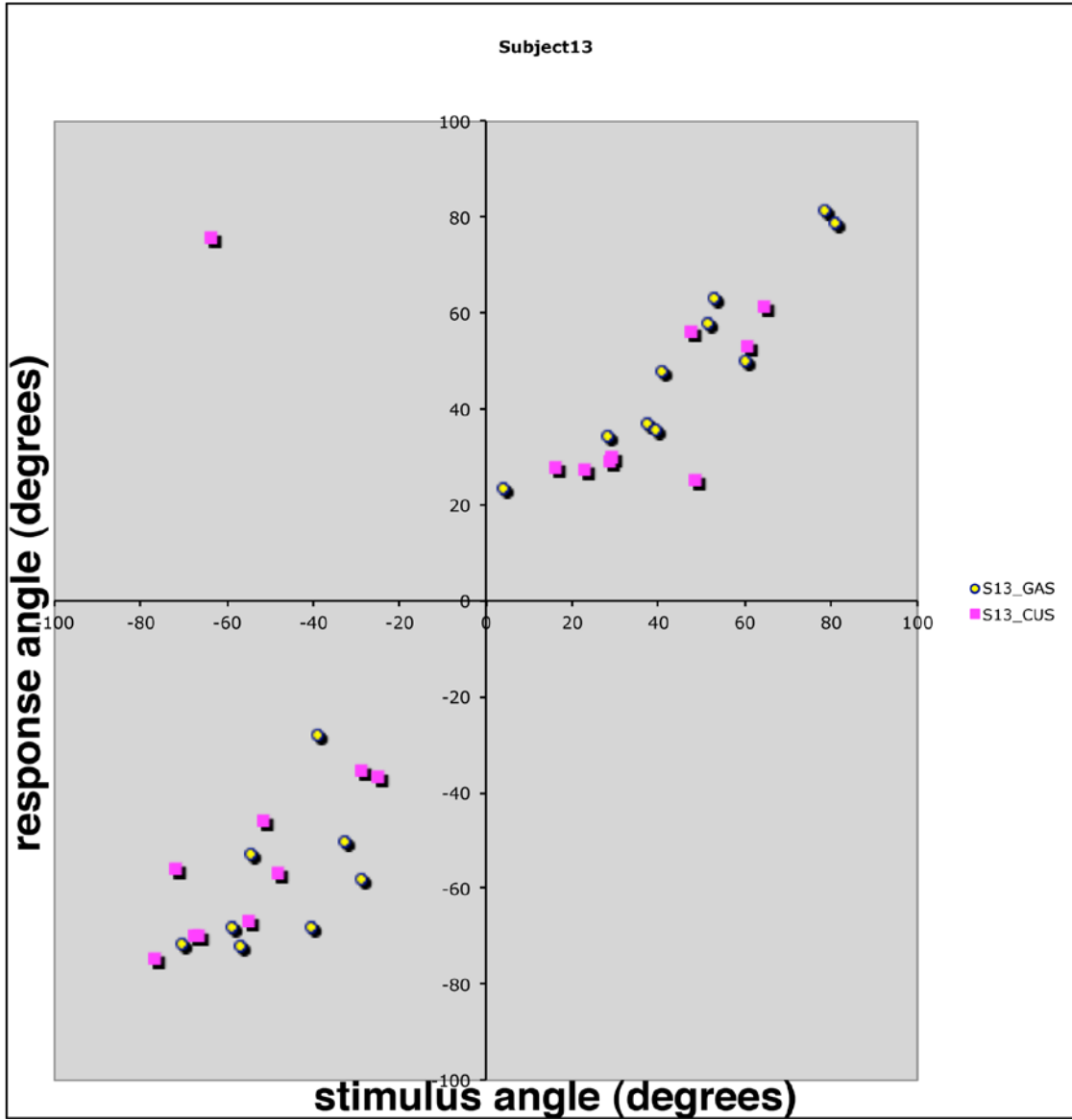


Figure A.13: Visualization data for subject 13.

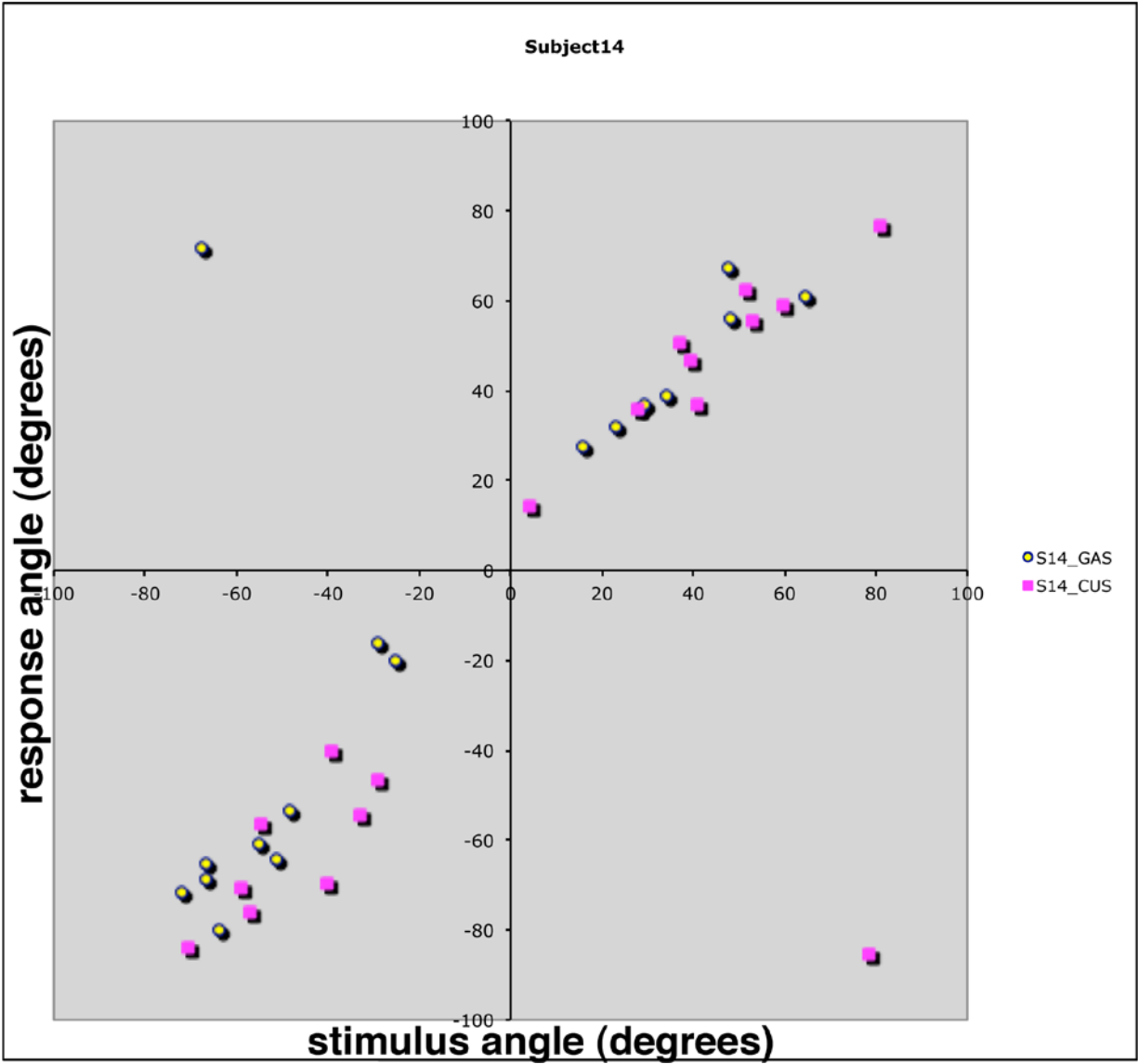


Figure A.14: Visualization data for subject 14.

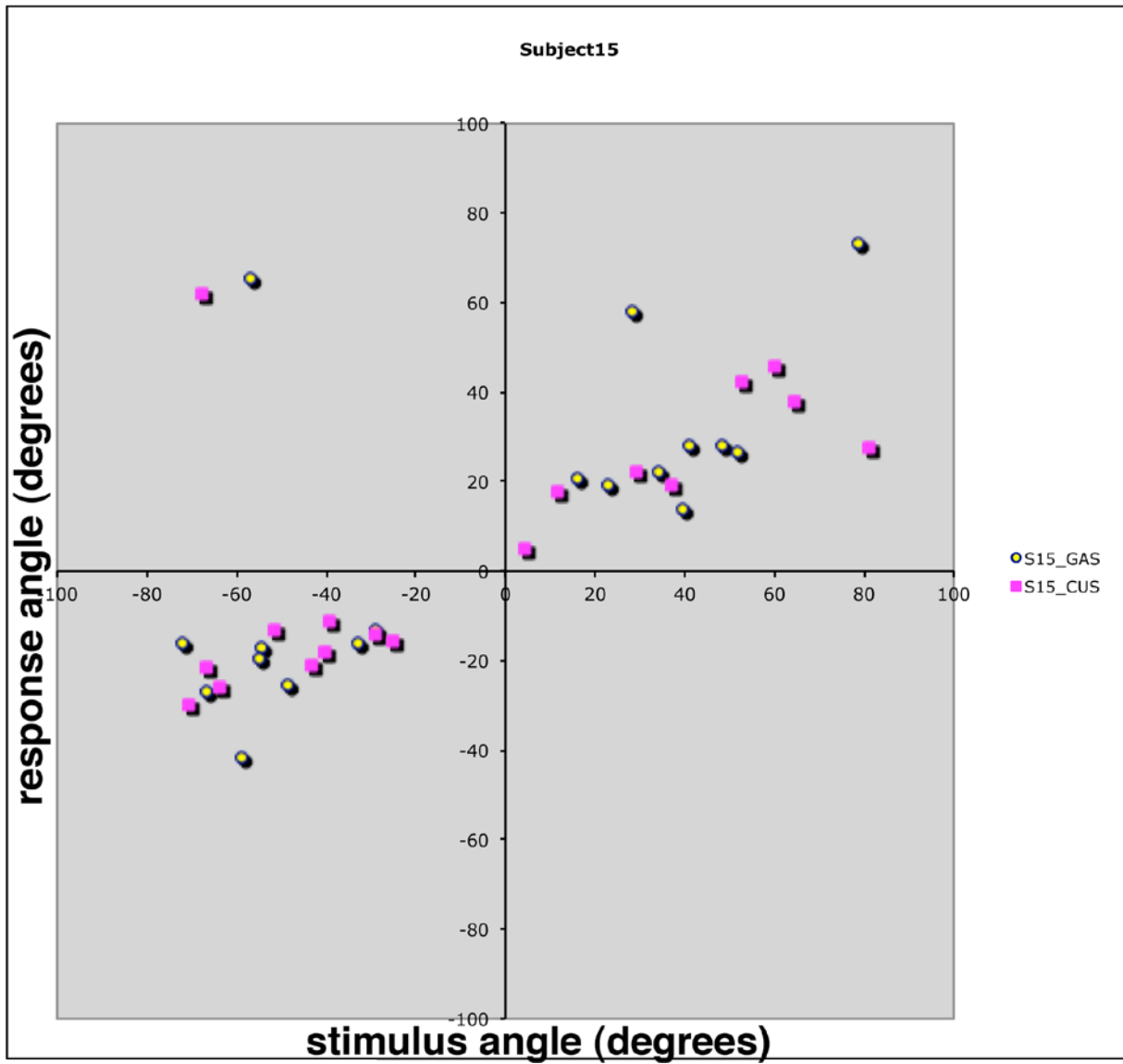


Figure A.15: Visualization data for subject 15.

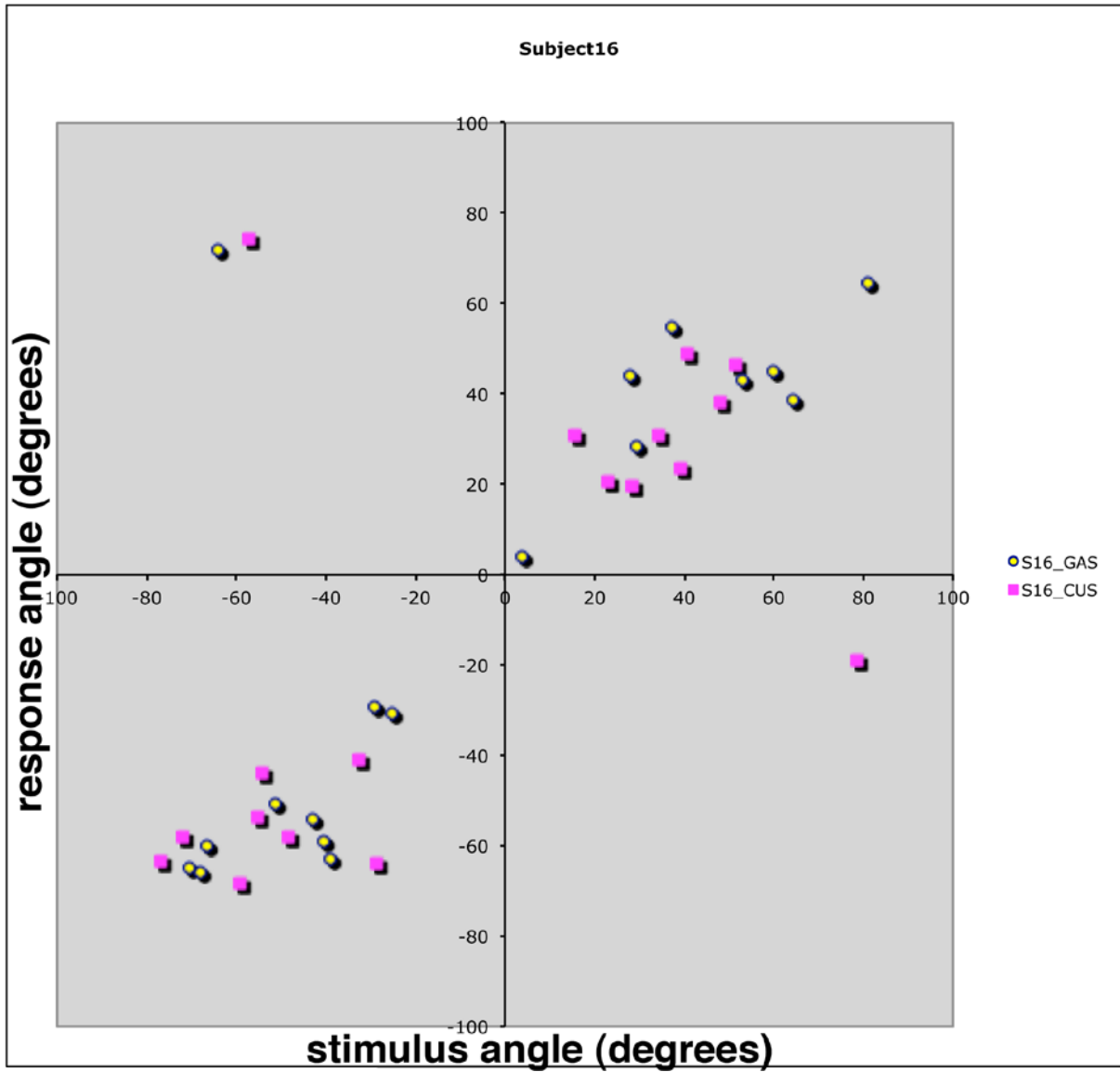


Figure A.16: Visualization data for subject 16.

APPENDIX B

RELEVANT SOFTWARE

Software has been included as a separate contribution in the electronic submission of this dissertation, in the file Code_ShuklaG_etd_2011.zip.

BIBLIOGRAPHY

- [1] Stetten, G, Chib, V. Overlaying Ultrasound Images on Direct Vision, *Journal of Ultrasound in Medicine*, vol. 20, no. 3, pp. 235-240, 2001.
- [2] Hofstein, S., "Ultrasonic Scope," US Patent no. 4,200,885, Apr. 1980.
- [3] Masamune, K., Masutani, Y., Nakajima, S., Sakuma, I., Dohi, T., Iseki, H., and Takakura, K. 2000. Three-Dimensional Slice Image Overlay System with Accurate Depth Perception for Surgery. *Proceedings of 3rd Int'l Conf. on Medical Image Computing and Computer-Assisted intervention* (October 11 - 14, 2000). S. L. Delp, A. M. DiGioia, and B. Jaramaz, Eds. *Lecture Notes In Computer Science*, vol. 1935. Springer-Verlag, London, 395-402.
- [4] Masamune, K., Fichtinger, G., Deguet, A., Matsuka, D., and Taylor, R. 2002. An Image Overlay System with Enhanced Reality for Percutaneous Therapy Performed Inside CT Scanner. In *Proceedings of the 5th international Conference on Medical Image Computing and Computer-Assisted intervention-Part II* (September 25 - 28, 2002). T. Dohi and R. Kikins, Eds. *Lecture Notes In Computer Science*, vol. 2489. Springer-Verlag, London, 77-84.
- [5] Masamune, K., Sato, I., Liao, H., and Dohi, T. 2008. Non-metal Slice Image Overlay Display System Used Inside the Open Type MRI. In *Proceedings of the 4th international Workshop on Medical Imaging and Augmented Reality* (Tokyo, Japan, August 01 - 02, 2008). T. Dohi, I. Sakuma, and H. Liao, Eds. *Lecture Notes In Computer Science*, vol. 5128. Springer-Verlag, Berlin, Heidelberg, 385-392.
- [6] Chang, W., Amesur, N., Klatzky, R., Zajko, A. , and Stetten, G., *Vascular Access: Comparison of US Guidance with the Sonic Flashlight and Conventional US in Phantoms. Radiology*, vol. 241, pp. 771-779, Dec. 2006.
- [7] Wu, B., Klatzky, R., Shelton, D., Stetten, G., *Psychophysical Evaluation of In-situ Ultrasound Visualization. IEEE Transactions on Visualization and Computer Graphics (TVCG), Special Issue on Haptics, Virtual and Augmented Reality*, vol. 11, no. 6, pp. 684-693, November/December 2005.
- [8] Klatzky, R., Wu, B., Shelton, D., Stetten, G., *Learning to perform actions in near space under perceptual vs. cognitive control. ACM Transactions on Applied Perception*. Vol. 5, No. 1, Article 1, January 2008.

- [9] Wu, B., Klatzky, R., Shelton, D., Stetten, G., Mental concatenation of perceptually and cognitively specified depth to represent locations in near space. *Experimental Brain Research*, Vol. 184, No. 3, pp. 295-305, 2008.
- [10] Klatzky, R., Wu, B., Stetten, G., Spatial representations from perception and cognitive mediation: The case of ultrasound. *Current Directions in Psychological Science*, December 2008, Vol. 17, No. 6., pp. 359-364.
- [11] B. Wu, R. Klatzky, G. Stetten, Visualizing 3D objects from 2D cross sectional images displayed in-situ versus ex-situ. *Journal of Experimental Psychology Applied*, 2010 Mar; 16(1):45-59.
- [12] W. Lorensen, H. Cline: *Marching Cubes: A high resolution 3D surface construction algorithm*. *Computer Graphics*, Vol. 21, Nr. 4, July 1987.
- [13] Marc Levoy, *Display of Surfaces from Volume Data*, *IEEE Computer Graphics and Applications*, v.8 n.3, p.29-37, May 1988.
- [14] JM Wolf, KR Kluender, DM Levi, LM Bartoshuk, RS Herz, RL Klatzky, SJ Lederman. *Sensation and Perception*. 1st ed. Sunderland, MA: Sinauer Associates, Inc., 2006.
- [15] Mike Blackwell, Constantinos Nikou, Anthony M DiGioia, Takeo Kanade, An Image Overlay system for medical data visualization, *Medical Image Analysis*, Volume 4, Issue 1, March 2000, Pages 67-72.
- [16] Ware, C. and Osborne, S. 1990. Exploration and virtual camera control in virtual three dimensional environments. In *Proceedings of the 1990 Symposium on interactive 3D Graphics (Snowbird, Utah, United States)*. SI3D '90. ACM, New York, NY, 175-183.
- [17] Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. 1994. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating interdependence (Boston, Massachusetts, United States, April 24 - 28, 1994)*. B. Adelson, S. Dumais, and J. Olson, Eds. CHI '94. ACM, New York, NY, 452-458.
- [18] Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (The Hague, The Netherlands, April 01 - 06, 2000)*. CHI '00. ACM, New York, NY, 526-531.
- [19] Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R., and Hart, J.C. The CAVE, Audio-Visual Experience Automatic Virtual Environment. *Communications of the ACM*, June 1992, pp 67-72.
- [20] Browning, D.R., Cruz-Neira, C., Sandin, D.J., DeFanti T.A. *Projection-Based Virtual Environments and Disability*. *Virtual Reality Conference*, 1994.
- [21] Looser, J., Billingham, M., and Cockburn, A. 2004. Through the looking glass: the use of lenses as an interface tool for Augmented Reality interfaces. In *Proceedings of the 2nd*

- international Conference on Computer Graphics and interactive Techniques in Australasia and South East Asia (Singapore, June 15 - 18, 2004). S. N. Spencer, Ed. GRAPHITE '04. ACM, New York, NY, 204-211.
- [22] Kalkofen D., Mendez E., Schmalstieg D., Interactive Focus and Context Visualization for Augmented Reality. In Proceedings of 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2007), Nara, Japan, November 2007 pp. 191-200, 2007.
- [23] State, A., Livingston, M. A., Garrett, W. F., Hirota, G., Whitton, M. C., Pisano, E. D., and Fuchs, H. 1996. Technologies for augmented reality systems: realizing ultrasound-guided needle biopsies. In Proceedings of the 23rd Annual Conference on Computer Graphics and interactive Techniques SIGGRAPH '96. ACM, New York, NY, 439-446.
- [24] Sauer F., Khamene A., Bascle B., Schimmang L., Wenzel F., Vogt S. Augmented Reality Visualization of Ultrasound Images: System Description, Calibration, and Features. pp. 30, IEEE and ACM International Symposium on Augmented Reality (ISAR'01), 2001.
- [25] Ito, T., Cassinelli, A., Komuro, T., and Ishikawa, M. 3D Object Representation Using a Tangible Screen (Volume Slicing Display). Proceedings of the 7th SICE System Integration Division Annual Conference SI2006, Sapporo, Japan, Dec. 14 – 17, 2006.
- [26] Kenneth T. Horlander, David M. Mannino, and Kenneth V. Leeper. Pulmonary Embolism Mortality in the United States, 1979-1998: An Analysis Using Multiple-Cause Mortality Data. *Arch Intern Med*, Jul 2003; 163: 1711 - 1717.
- [27] Silverstein MD, Heit JA, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ 3rd. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch Intern Med*. Mar 23 1998;158(6):585-93.
- [28] Kline JA, Runyon MS. Pulmonary embolism and deep venous thrombosis. In: Marx JA, Hockenberger RS, Walls RM, eds. *Rosen's Emergency Medicine Concepts and Clinical Practice*. Vol 2. 6th ed. 1368-1382.
- [29] Heit JA. The epidemiology of venous thromboembolism in the community. *Arterioscler Thromb Vasc Biol*. Mar 2008;28(3):370-2.
- [30] Dzung L Pham, Chenyang Xu, Jerry L Prince. Current Methods in Medical Image Segmentation. *Annual Review of Biomedical Engineering*, August 2000; 2:315-337.
- [31] A Fenster, B Chiu. Evaluation of Segmentation Algorithms for Medical Imaging. Proceedings of 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Shanghai, China, Sept 1-4, 2005: 7186-7189.
- [32] Chuan Zhou, Heang-Ping Chan, Berkman Sahiner, Lubomir M. Hadjiiski, Aamer Chughtai, Smita Patel, Jun Wei, Jun Ge, Philip N. Cascade, and Ella A. Kazerooni. Automatic multiscale enhancement and segmentation of pulmonary vessels in CT pulmonary angiography images for CAD applications. *Med. Phys*. December 2007, 34:4567-4577.

- [33] S. R. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE Trans. Med. Imaging* 21, 61–75, 2002.
- [34] Fairfield, J. Toboggan contrast enhancement for contrast segmentation. Proceedings of the 10th International Conference on Pattern Recognition. Jun 1990, pp 712-716.
- [35] J. Liang, M. Wolf, and M. Salganicoff, "A fast toboggan-based method for automatic detection and segmentation of pulmonary embolism in ct angiography," in *The Eighth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2005)*.
- [36] Guo Li, Jianhua Wu, Pian Zhao-yu, Wang Kun. Improved Toboggan Segmentation Algorithm for Magnetic Resonance Images. 2nd IEEE Conference on Industrial Electronics and Applications, May 23-25, 2007, pp. 2504-2507.
- [37] Abolmaesumi, P.; Sirouspour, M.R.; Salcudean, S.E., "Real-time extraction of carotid artery contours from ultrasound images," *Computer-Based Medical Systems, 2000. CBMS 2000. Proceedings. 13th IEEE Symposium on* , vol., no., pp.181-186, 2000.
- [38] Wang D, Klatzky R, Wu B, Weller G, Sampson A, Stetten G. Fully automated common carotid artery and internal jugular vein identification and tracking using B-mode ultrasound. *IEEE Transactions on Biomedical Engineering* (in press).
- [39] C.A. Cois, *Variable Scale Statistics for Cardiac Segmentation and Shape Analysis*, Ph.D. Dissertation, Department of Bioengineering, University of Pittsburgh, advisor G. Stetten, October, 2007.
- [40] C.A. Cois, K. Rockot, J. Galeotti, R. Tamburo, D. Gottlieb, J. Mayer, A. Powell, M. Sacks, G. Stetten, Automated segmentation of the right heart using an optimized shells and spheres algorithm, *International Symposium on Biomedical Imaging (ISBI)*, April 2007.
- [41] C.A. Cois, K. Rockot, J. Galeotti, R. Tamburo, G. Stetten, *Shells and Spheres: A Framework for Variable Scale Statistical Image Analysis*, CMU Robotics Tech Report #CMU-RI-TR-04-19, April, 19, 2006.
- [42] C.A. Cois, R. Tamburo, J. Galeotti, M. Sacks, G. Stetten, *Shells and Spheres: An n-Dimensional Framework for Medial-Based Image Segmentation*, *International Journal of Biomedical Imaging*, special issue on *Mathematical Methods for Images and Surfaces*, 2010.
- [43] V.R. Shivaprabhu, *Automated Method for N-Dimensional Shape Detection Based On Medial Image Features*, Masters' Thesis, Department of Electrical Engineering, University of Pittsburgh, advisor G. Stetten, December 2010.
- [44] G. Stetten, S. Horvath, J. Galeotti, G. Shukla, B. Chapman, Image segmentation using the student's t-test on adjacent spherical populations of pixels. *SPIE Medical Imaging*, paper #7623-125, poster presentation 2/14-15/2010.

- [45] DZ Reinstein, RH Silverman, HFS Sutton, DJ Coleman. Very High-Frequency Ultrasound Corneal Analysis Identifies Anatomic Correlates of Optical Complications of Lamellar Refractive Surgery. *Ophthalmology*; March 1999, 106:474-482.
- [46] D Lee, MT van Holsbeeck, PK Janevski, DL Ganos, DM Ditmars, VB Darian. Diagnosis of Carpal Tunnel Syndrome: Ultrasound Versus Electromyography. *Radiologic Clinics of North America*. July 1999: 37:859-872.
- [47] U.S. Patent no. 6,599,247, G. Stetten, System and Method for Location-Merging of Real-Time Tomographic Slice Images with Human Vision, filed Oct. 11, 2000, issued, July 29, 2003.
- [48] G Shukla, B Wu, D Schwartzman, GD Stetten, The Sonic Penlight for Guidance of Superficial Subdermal Access,” *Medical Image Computing and Computer-Assisted Intervention Conference*, Oct. 26-29, 2005, Palm Springs, CA. (Abstract + Poster)
- [49] J. Galeotti, A. Sajjad, B. Wang, L. Kagemann, G. Shukla, M. Siegel, B. Wu, R. Klatzky, G. Wollstein, J. Schuman, G. Stetten, The OCT penlight: In-situ image guidance for microsurgery, *SPIE Medical Imaging*, paper #7625-1, oral presentation 2/14/2010.
- [50] W. Chang, N. Amesur, D. Wang, A. Zajko, G. Stetten First Clinical Trial of the Sonic Flashlight - Guiding Placement of Peripherally Inserted Central Catheters, 2005 meeting of the Radiological Society of North America, November 2005, Chicago, Illinois.
- [51] N. Amesur, D. Wang, W. Chang, D. Weiser, R. Klatzky, G. Shukla, G. Stetten, Peripherally Inserted Central Catheter Placement by Experienced Interventional Radiologist using the Sonic Flashlight, *Journal of Vascular and Interventional Radiology*, 2009; 20:1380–1383.
- [52] D. Wang, N. Amesur, G. Shukla, A. Bayless, D. Weiser, A. Scharl, D. Mockel, C. Banks, B. Mandella, R. Klatzky, G. Stetten, Peripherally Inserted Central Catheter Placement with the Sonic Flashlight: Initial Clinical Trial by Nurses, *Journal of Ultrasound in Medicine*, 28:651-656, 2009.
- [53] M. Chen, G. Shukla, S. Jackman, A. Tsao, M. Smaldone, M. Ost, G. Stetten, T. Averch, Real-time tomographic reflection in facilitating percutaneous access to the renal collecting system, *Journal of Endourology* (in press).
- [54] D. Shelton, Virtual Tomographic Reflection for Psychophysical Analysis of the Sonic Flashlight, Ph.D. Dissertation, Robotics Institute, Carnegie Mellon University, advisor G. Stetten, March 2007.
- [55] G Shukla, B Wang, J Galeotti, R Klatzky, B Wu, B Unger, D Shelton, B Chapman, and G Stetten, A Movable Tomographic Display for 3D Medical Images, *Proceedings of the Augmented Environments for Medical Imaging including Augmented Reality in Computer-Aided Surgery (AMI ARCS)*, at *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, p. 69-76, London, September 24, 2009.

- [56] A Watt, *Fundamentals of Three-Dimensional Computer Graphics*. Reading, MA: Addison-Wesley, 1990.
- [57] G Shukla, R Klatzky, B Wu, B Wang, J Galeotti, G Stetten. Psychophysical evaluation of a movable tracked tomographic display. *IEEE Trans Visualization and Computer Graphics*, (submitted).
- [58] Van Beurden, M.H.P.H., van Hoey, G., Hatzakis, H., & IJsselsteijn, W.A. Stereoscopic displays in medical domains: a review of perception and performance effects, *Human Vision and Electronic Imaging XIV*, Proceedings of the SPIE, 7240, 72400A-72400A-15, 2009.
- [59] Hubbard, R.J., Hancock, D.J., Moore, C.J. "Autostereoscopic display for radiotherapy planning," *Proceedings SPIE*, 3012, 16-27, 1999.
- [60] Moll, T.H., Turjman, F., Picard, C., Bres, J.P., Amiel, M. Depth separation in ten observers with a new stereoscopic X-ray acquisition system, *European Radiology*, 7(8), 1343-1347, 1997.
- [61] Doi, K., Duda, E.E. Detectability of depth information by use of magnification stereoscopic technique in cerebral angiography, *Radiology*, 146, 91-95, 1983.
- [62] Rosenbaum, A.E., Huda, W., Lieberman, K.A., Caruso, R.D. Binocular three-dimensional perception through stereoscopic generation from rotating images, *Academic Radiology*, 7(1), 21-26, 2000.
- [63] Kickuth, R., Hartung, G., Laufer, U., Gruening, C., Stueckle, C., Liermann, D. Stereoscopic 3D CT vs. standard 3D CT in the classification of acetabular fractures: An experimental study, *British Journal of Radiology*, 75(893), 422-427, 2002.
- [64] Hernandez, A., Basset, O., Bremond, A., Magnin, I.E. Stereoscopic visualization of three-dimensional ultrasonic data applied to breast tumours, *European Journal of Ultrasound*, 8(1), 51-65, 1998.
- [65] Hsu, J., Babbs, C.F., Chelberg, D.M., Pizlo, Z. Study of the effectiveness of stereo imaging with applications in mammography, *Proceedings of SPIE*, 1913, 154-165, 1993.