2015

Human factors and performance considerations of visual spatial skills in medical context tasks

Marisol Martinez Escobar
Iowa State University

Human Factors and performance considerations of visual spatial skills in medical context tasks

by

Marisol Martinez Escobar

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-majors: Mechanical Engineering, Human Computer Interaction

Program of Study Committee:
Eliot Winer, Major Professor
Jonathan Claussen
Matt Frank
Stephen Gilbert
Jim Oliver

Iowa State University

Ames, Iowa

2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

ABSTRACT

In the medical field, stereoscopic applications are present in diagnosis, pre-operative planning, minimally invasive surgery, instruction, and training. The use of stereoscopic applications has afforded new ways to interact with patient data, such as immersive virtual environments. This increased usage of stereoscopic applications also raises many basic research questions on human perception and performance.

Current studies show mixed results on the benefits of stereoscopic applications with regards to general performance. The benefits depend on the specific task as well as the application domain. The work presented here attempts to answer the general question: How would adding the stereopsis depth cue affect the performance of visual spatial tasks in a medical context? Visual spatial tasks are needed in medicine to understand the relationships between shapes and organs for a variety of activities in patient diagnosis and treatment.

The general research question was decomposed into specific hypotheses and three studies were conducted to study them. These studies measured performance of a visual spatial computer task using medical imaging data. Participants assessed the relative positions of three different objects located inside a 3D volumetric representation of a patient's anatomy. The first study consisted of static views and recognition of the position of color objects. The second study consisted of static views using gray objects. The third study consisted of animated views of color objects. In all three studies the task was basically the same: To select which of two objects was closest to a reference object. In all three studies participants were first and second year medical students.

Thirty-four participants completed the first study. The results of this study showed some emerging patterns in which the stereoscopic display condition had a positive benefit on performance. The stereoscopic condition had a positive effect on performance for the most difficult cases but did not yield higher results under every case and condition.

The second study, completed by 44 participants, showed the stereoscopic condition had a positive benefit on performance in 20 out of the 40 tasks completed. These 40 tasks were divided into four cases, with varying degrees of difficulty, depending on the distances between the objects being judged (i.e. cylinders in this study). At distances between 5-15 mm, the stereoscopic condition yielded statistically significant higher performance. At other distance ranges, while stereopsis showed improvement it was not statistically significant.

Thirty-one participants completed the third study. These participants completed a visual spatial task with the addition of an animation to the volume. This allowed the representation to be viewed from multiple angles before the task was completed. Overall the stereoscopic condition had a benefit in performance over the monoscopic condition. As in the previous studies tasks that had the objects between 5 – 15 mm apart had higher performance in the stereoscopic condition. Females performance in the stereoscopic condition was higher and statistically significant than for the monoscopic condition. Participants over 25 years also had a statistically significant higher performance under the stereoscopic condition. It was also observed that the stereoscopic condition did not outperform the monoscopic one in every condition.

The results of these studies show that, in general, stereopsis has a positive benefit in performance for visual spatial tasks in medical contexts. This benefit certainly has a relationship with the difficulty of the task as well as age and gender. These initial insights are a step into further work to help generate design guidelines when developing stereoscopic applications.

CHAPTER 1. INTRODUCTION

Understanding depth information is an essential skill used to make sense of the world around us. Depth information helps humans recognize shapes and objects, understand position of objects relative to each other, understand the position of objects relative to the self, manipulate real and digital objects, and grasp objects. Many of these tasks are tied to human survival, such as estimating the position of a dangerous animal, or grasping fruit to eat. Because of the importance of depth information for humans, the Human Visual System (HVS) has redundant cues working together to get this information. Some of these depth cues are obtained with the use of only one eye, monocular cues. Some of the cues can only be obtained with both eyes, binocular cues.

It has been shown that in real life that performance is degraded in some tasks when binocular cues are missing [1-3]. Melmoth [2] conducted a reach and grasp study, 16 participants using goggles were asked to grasp and object, the goggles simulated binocular and monocular vision. Participants were asked to reach an object in front of them and place it on the table, and to move as naturally, quickly and accurately as possible. The performance of participants with binocular condition was higher than for the monocular condition.

Even though binocular cues afford better performance on some tasks, and have potential to improve performance on many other tasks, conventional displays with only monocular cues are the norm. Displays with binocular depth cues have been largely used in the entertainment industry as a way to improve the look and feel of movies and games.

But there is potential to use displays with binocular depth cues in other areas such as medicine [4], military [5], and engineering [6].

In medicine, stereopsis, one of the most sensitive binocular depth cues, has been used to understand shapes and objects, assess the relative position of objects, and in grasping objects [1, 4, 7]. Given that medical errors can lead to death or permanent injury at a rate of 200,000 lives per year in the US alone [8], new tools that improve medical performance are needed but these tools must first address current limitations and show benefits before they are adopted into the medical field.

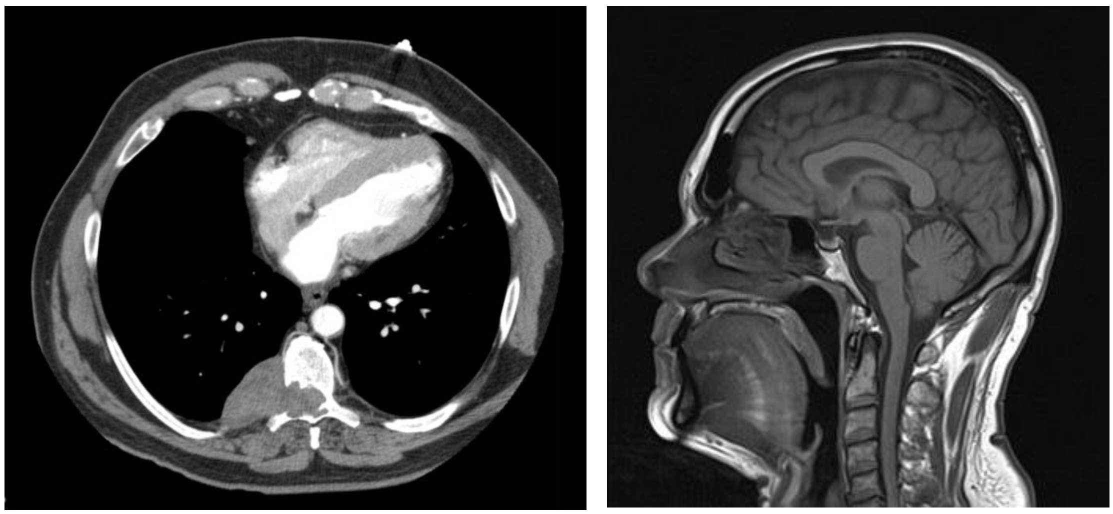
1.1 Medical Visualization History on Innovation

The medical field has a rich history of innovation, as its always looking to improve patient care. Because of current problems to be solved and their record with adopting new technologies, medicine can be an area where stereoscopic applications have a positive benefit. An example of innovation in the medical field is found in the history of the X-ray. Willhelm Rontgen discovered the X-rays in 1895, the image of the first X-ray is represented in Figure 1. Rontgen published an article on them on a physics and medical journal. Within eight days of the publication of the article on X-ray imaging, the X-ray was used to remove an industrial sewing needle from a woman's hand [9], a few months later an the X-ray was used to successfully remove a bullet [10]. The X-ray afforded a new way to see inside the body, not possible before.

Figure 1. First X-ray taken in 1895

The medical imaging community continued their innovation with the advancement of imaging techniques, such as Computed Tomography (CT) [11] and Magnetic Resonance Imaging (MRI) [12], Figure 2 shows an example of a CT and a MRI scan of a patient’s abdomen and head respectively. In 1972, Sir Geoffrey Hounsfield introduced the first CT machine. CT uses X-ray technology to produce a series of images, affording additional views not possible with a single X-ray. Paul Lauterbur and Peter Mansfield introduced the MRI in 1977. MRI machines produce body images using a magnetic field. CT and MRI machines can generate images in approximately 30 milliseconds.

Integrating and understanding the vast amount of medical images generated from CT or MRI techniques is a challenge, Figure 3 shows multiple images generated from a CT or MRI scan of a patient’s head. Volume rendering techniques have been developed as a way to facilitate the understanding of large amounts of data. As its name implies volume rendering techniques combine the image information of 2D medical images, called slices, and generate a 3D representation of the data, Figure 4 shows how 2D images of a patient are composed into a 3D representation. Volume rendering techniques are out of scope of this paper but described in detail in [13-16].



a) CT scan of the abdomen b) MRI scan of the head

Figure 2. Examples of CT scans and MRI scans of body parts.

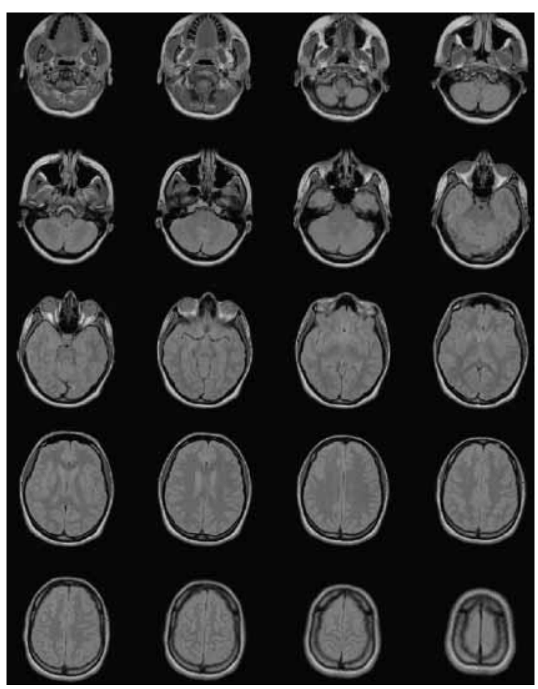


Figure 3. Multiple scans obtained from a CT study.

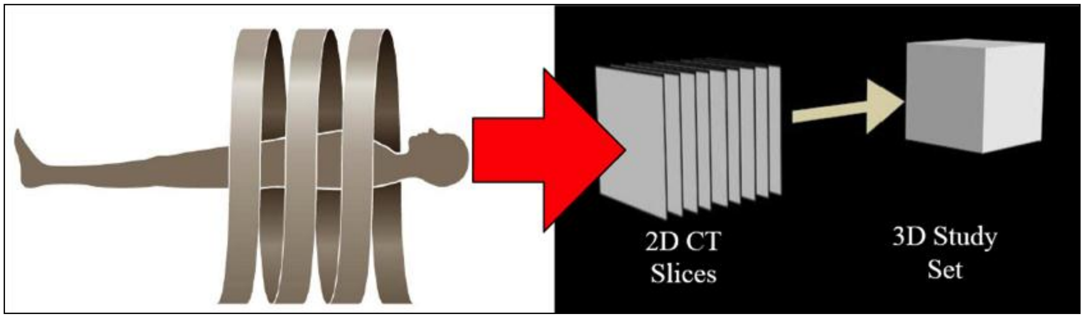


Figure 4. Volume rendering process from 2D slices.

Software implementations that are able to process large amounts of data have been designed to visualize medical images. These applications display the information through graphical user interfaces and allow the user to interact with the medical images. Examples of these software applications include Osirix [17], Volview [18], and BodyViz [19]. BodyViz desktop software allows for the visualization of MRI/CT slices into a volume. Using a controller the user can interact with the volume to change the rotation, the color, the tissues to be displayed, the position, zoom in into the data, and other types of transformations. Figure 5 shows the interface of BodyViz with medical data of an abdomen, these images are colored to highlight body features of interest.

Figure 5. BodyViz Software for the visualization of medical data [19].

1.2 Motivation

Currently most displays use traditional depth cues, which are not as robust as binocular depth cues. Stereopsis is one of the most important binocular depth cues, but the inclusion of stereoscopic applications for performance related tasks has been low. In the medical field, where errors can cost a patient's live, stereoscopic applications could potentially offer performance benefits not available in traditional displays. Stereoscopic applications have not been widely adopted in the medical field because of three main factors: 1) negative physiological displays associated with stereoscopic displays, 2) additional costs in hardware, and 3) the unclear understanding of the advantages of stereoscopic displays over traditional monoscopic displays.

The first contributing factor against the adoption of stereoscopic displays is the association of stereoscopic displays and physiological effects. It is estimated that around two to three million persons in the US cannot see stereopsis and when presented with stereoscopic applications they experience negative reactions, such as headaches, blurred vision, and dizziness when using stereoscopic displays [20-28].

The second contributing factor to the adoption of stereoscopic displays is cost. Stereoscopic systems require additional equipment such as glasses, special displays, and even specialized tracking systems. In some cases the computing requirements are more demanding than regular displays [4]. However, processing power continues to increase while simultaneously decreasing in cost, such that a stereoscopic display can be acquired for little more than a standard desktop computer.

The third contributing factor slowing the adoption of stereoscopic displays, and the focus of this work, is whether stereopsis increases performance of medical tasks. Of course, there is a wide range of tasks that occur daily in the practice of medicine, so answering this general question is extremely challenging. Many of the studies that compare monoscopic and stereoscopic applications have yielded positive, mixed, or even negative results [4], these results are dependent of domain and tasks. For example, a stereoscopic applications survey [29] showed that out of 12 studies on performance of navigational tasks, seven out of the 12 studies showed no benefit or negative benefit of the stereoscopic condition. However, none of these studies were conducted in a medical context or environment. In contrast there were 26 studies on the performance of finding and identifying objects using stereoscopic applications, and 17 of 26 studies showed a benefit of stereopsis, and nine of the 26 studies showed no benefit or negative benefits of stereoscopic applications.

The vast majority of the current studies of performance on medical applications center around the manipulation of objects, with a lack of work of other tasks such as navigation and relative judgment of position and distances [29]. Relative judgment of position and distances of objects is an important skill in medicine to understand anatomy and how organs and tissues are positioned inside the body [30, 31]. This is a skill needed for diagnosis, planning, surgery, and overall treatment of patients.

The focus of this work is to conduct human factors studies on the performance of stereoscopic applications for relative position judgments in medical contexts. By understanding the relationship between performance, and human factors, these results and further work could eventually help in the formation of guidelines for designing stereoscopic applications in medicine. Understanding what are the benefits of stereoscopic applications is extremely important for the medical community as it has potential of reducing errors and improve patient care.

1.3 Dissertation Organization

This dissertation is organized as follows: Chapter 2 presents the background information on HVS and depth cues, as well as examples of stereoscopic applications in medicine. Chapter 3 presents the overall methodology used in the three studies. Chapter 4 presents a published conference paper of the results of the first study. Chapter 5 presents additional analysis of the first study that were not in the conference paper due to length requirements. Chapter 6 presents a submitted journal paper of the results of the second study. Chapter 7 presents additional analysis of the second study. Chapter 8 presents a second journal submitted with the results of the third study. And finally Chapter 9 presents conclusions and future work.

CHAPTER 2. BACKGROUND

2.1 Depth and the Human Visual System

In order to see depth the Human Visual System (HVS) uses a combination of monoscopic and binocular depth cues [32, 33]. Monocular depth cues can be represented in conventional displays, and they include perspective projection, occlusion, familiar size, object motion, accommodation, and motion parallax [4, 32]. Binocular depth cues include stereopsis, and convergence.

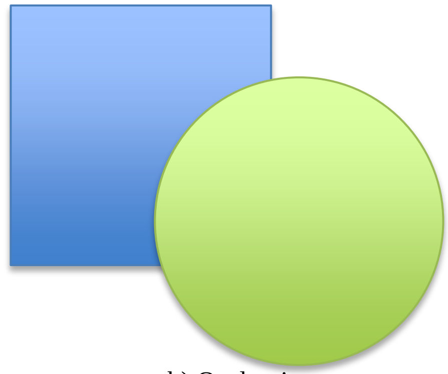
2.1.1 Monocular Depth Cues

Perspective projection is the display of a 3D scene into a flat plane, such as a conventional TV. This monocular cue makes use of sizes, lines, and textures. As in real life, objects that are closer to the observer are projected in bigger sizes than objects that are far away. Similarly, lines that are farther apart are closer to the observer, and lines that converge are farther away from the observer.

Occlusion is another monocular cue in which objects in the front block objects that are farther apart. One of the limitations of this cue is its strength. While blocking objects allows the observer to make judgments on order, farther away objects can be hidden and judgment on distances are not possible. For example, in Figure 6b the circle is closer to the observer because it's occluding the square.

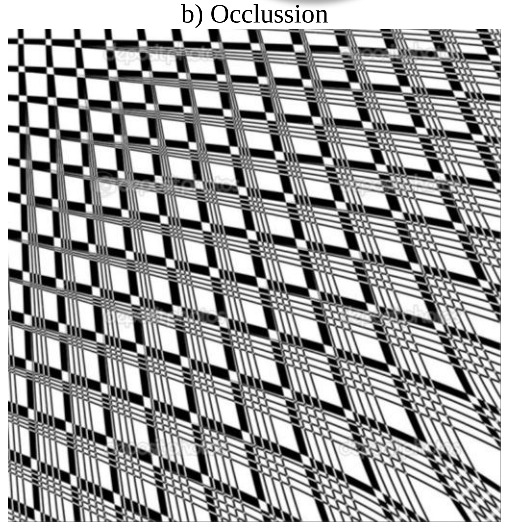
An example of perspective and occlusion cues is shown Figure 6a, there are several monocular cues that help in understanding object depth in the scene. Objects in the front, such as the people walking the street are occluding objects in the back. Taxis that are bigger in size are closer to the observer. Larger distances between buildings signify they are closer to the observer.

a) Perspective



b) Occlussion

c) Familiar Size



d) gradient

Figure 6. Examples of monocular depth cues.

“Familiar” size is another important monocular depth cue that uses a known object as a reference. A familiar object provides reference points for the sizes and depths of other objects in a scene. For example, in Figure 6c the man in the picture is closer to the observer as a person is typically not the same size as a tower. Using the tower as a reference object in the image, an observer can understand the appropriate sizes and positions of other objects in the scene.

Texture gradient is another monocular cue that provides depth information. Details on textured surfaces are seen more clearly when they are closer to an observer. For example in Figure 6d, the observer can see additional lines on the shapes that are closer. This, combined with the objects shape, provide the appearance of the top left of the image being farther away from the observer.

Object motion, commonly obtained by rotating an object around its axis, is a monocular cue that allows an observer to see an object from different angles and construct a 3D representation. Usually the observer is static while the object is in motion.

Motion parallax is the opposite scenario with an observer moving while the object, or scene remains static. The observer understands depth by the relative motion of the objects in the scene. Closer objects will move faster and longer distances than objects that are farther away. For example, when driving a car, trees in the scenery that are closer to the car seem to move much faster than the objects (e.g., mountains, hills) off in the distance.

Accommodation is another monocular depth cue. When trying to focus on objects the intraocular muscles of the eyes contract and send this information to the visual cortext to interpret depth information.

All, or varying combinations, of the preceding depth cues are interpreted and used by the human visual system to effectively create a depth map of scenes in real-time. While the monocular depth cues described above are important they are limited and not as sensitive as stereopsis [4, 34]. In some other cases, such as looking at objects close to the observer, monocular depth cues are not available and binocular depth cues are solely responsible for assessing depth.

2.1.2 Binocular Depth Cues

Binocular depth cues are obtained using both eyes, and include stereopsis and convergence. Convergence is the movement of the eyes to each other when trying to focus on an object. The effort of the extraocular muscles when focusing on an object is used by the visual cortext to assess depth interpretation. Convergence only works for distances less than approximately 10 m away [34].

Stereopsis is a binocular depth cue that consists of processing the disparity between unique images presented to each eye. As can be seen in Figure 7 human eyes are slightly separated horizontally and receive a unique image of the world. The brain fuses these left and right images and extracts depth information [21]. Stereopsis is a very sensitive depth cue, and is capable of differentiating objects millimeters away from one another [4]. This sensitivity offers potential to use stereopsis as a depth cue to improve performance on tasks where spatial awareness is critical.

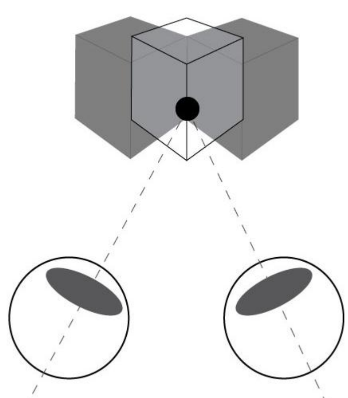


Figure 7. Illustration of how each eye obtains a slightly different image, when viewing an object.

2.2 Stereoscopic Technology

The basis of stereoscopic technology displays is to present a unique image to each eye. Different technologies have been developed providing multiple ways to do this using computer generated displays [34-36]. Three of the most popular stereoscopic techniques are:

- Anaglyph
- Active Stereo
- Passive Stereo

Anaglyph uses two slightly different superimposed color images on a single display. Using special glasses with appropriate color “filters”, each eye receives a unique image, creating the stereopsis effect. Anaglyph technologies have been used in the entertainment industry in the past, typically using red-green filtering as shown in Figure 8. The advantages of anaglyph technology is that it uses conventional displays to create stereoscopic applications and is very low cost. However, because two images are presented on the same display at the same time, the images suffer degradation and generally poor visual quality. The resolution of the images is halved because both images have to share the image resolution the display [34]. The filtering requires certain colors to be removed from the images causing the overall scene to look somewhat unnatural or even incorrect when perceived by a human observer.

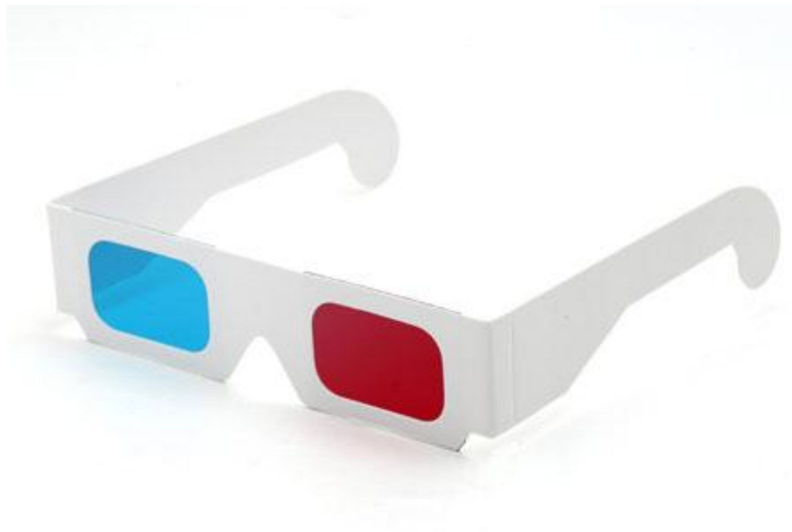


Figure 8. Anaglyph lenses.

In active stereo systems a suitable display shows an image for one eye and subsequently the image for the other eye. The system does this so rapidly that the brain is capable of fusing the images together to create the stereopsis effect. Active shutter technologies require glasses that will close in front of one eye, when the image for the other eye is being shown on a display. This typically occurs at a rate of 30-120 Hz. Active stereo applications can be used without affecting the resolution of the images, but displays or projectors capable of refreshing (i.e. showing the images on the screen) fast enough to provide at least 15 frames per second is required (30 or higher is preferred). While active stereo offers the best VR experience, the hardware needed to produce and view the images is the most expensive.

Passive stereo technologies again use filtering, but with more sophistication than anaglyph. Images on a display may be shown at the same time or in sequence. The glasses used will incorporate filtering, such as polarization, to remove one set of images for each eye. For example, the images for the left eye may be polarized in one way and those for the right in another. There are several schemes commonly used including side by side, and top/bottom. As their names suggest, side by side draws next to each other and top/bottom draws one over the other. Each type of polarizing glass is different depending on the polarization scheme [34]. Polarization is the most widely used for consumer VR viewing such as movies and high-definition televisions. This is due to the better than average quality at reasonable cost.

There are other stereo technologies that are adaptations of the ones described above or hybrids between the categories. For example, autostereoscopic displays do not require glasses to create a stereopsis effect. Autostereoscopic displays superimpose the right and left images in the display and use a parallax barrier to direct the light of each image to the appropriate eye. Autostereoscopic displays only work when the user is located at extremely precise positions relative to the display, otherwise the user will get light from the incorrect image to both eyes [34].

2.3 Stereoscopic Applications in Medicine

Current stereoscopic medical applications can be divided into four areas: diagnosis, pre-operative planning, minimally invasive surgery, and teaching/training [4, 37].

In diagnosis, evaluating shapes, sizes, and relative positions identifies abnormalities of tissues and organs. Stereopsis can benefit diagnosis with shape recognition, understanding relative positions of objects, and separating the object of interest from the background image. Several stereoscopic applications have been developed for diagnostic purposes, including applications in ophthalmology [19] and mammography [38]. The studies conducted on performance of stereoscopic applications in medicine have yielded mixed results. Kickuth et al. [39] conducted a study to classify fractures, participants classified the fractures in both the monoscopic and the stereoscopic condition. They found no benefits in using stereoscopic displays for diagnosing fractures. Getty et al. [38] conducted a study on diagnosing breast lesions between standard 2D and stereoscopic 3D, an example of digital mammograph used in the study is shown Figure 9. Five users completed the study, it was found that the stereoscopic condition had better performance, additionally the stereoscopic condition allow for additional identification of lesions, not possible with the 2D views.

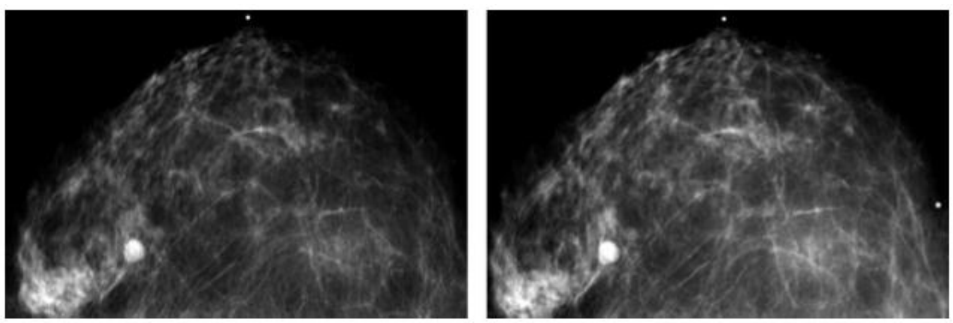


Figure 9. Stereoscopic pair of digital mammographs Getty [38].

Pre-operative planning is used to find the best path for surgery and design the procedure to the best of the surgeon's ability using digital medical data obtained from the patient prior to the procedure [40]. Pre-operative planning can improve surgical precision and reduce the time of an intervention as well as complications. Pre-operative planning skills involve identifying organs, understanding the relative positions between organs and surgical tools, and navigation of the surgical path. Several of these tasks can benefit from stereopsis. Several stereoscopic applications have been developed for pre-operative planning, but have offered mixed results when compared to monoscopic counterparts.

This makes it unclear if the benefits of stereopsis will translate to clinical applications [41].

Minimally invasive surgery (MIS) is another potential application for stereoscopic displays. MIS attempts to reduce trauma on the body by using smaller incisions and tools when compared to “open” procedures where significant portions of anatomy are exposed to the outside environment. In MIS several small incisions are created in the patient large enough to accommodate trocars, or narrow tubes that allow surgical instruments to be moved in and out. MIS has been proven to lower the recovery times for patients [42] and reduce errors in medicine. MIS is challenging for the surgeon because visual information is restricted [31]. The operative space cannot be seen directly, as it is only available through a camera (laparoscope) placed through one of the trocars. Traditionally the image obtained from the laparoscope is restricted to a 2D flat monitor and lacks many of the depth cues associated with open surgery. The lens of the laparoscope is fitted at an angle to expand the field of view and allows a surgeon to examine structures and objects in a specified region. However, this produces a non-intuitive perceptual view. For example, if the laparoscope of Figure 10 were to be rotated, a normal camera would need to be pitched up, but the angled laparoscope would need to be rotated 180 degrees. This is a different type of mapping that is challenging and reduces performance for novice users [43, 44].

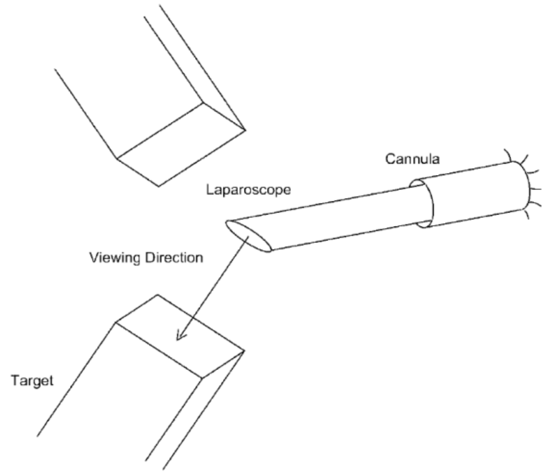


Figure 10. Angled lens of the camera inserted in Minimally Invasive Surgery [31].

Stereoscopic applications in MIS can help in navigation, estimating relationships between organs, recognizing shapes and sizes, and understanding the MIS mapping between cameras and displays. As with other stereopsis medical studies, those in MIS involving stereo have again yielded mixed results. Experienced surgeons prefer traditional MIS displays to stereoscopic displays [45, 46]. It is possible that experienced surgeons rely heavily on the use of monocular depth cues to understand depth without the need of stereopsis due to experience with procedures and surgical techniques in general. However, novice users have demonstrated benefit from stereopsis in MIS tasks.

Another medical application widely performed is anatomy training. Every doctor from surgeons to internists takes anatomy courses in medical school. Recently, computer simulations and visualizations have been used to facilitate this training. These simulations offer an alternative to cadavers, which are expensive, do not offer living tissue for study, and only offer limited study of anatomical abnormalities that will be encountered. Stereoscopic applications could aid in training by providing improved spatial representaiton of anatomical features. Several studies show an advantage of stereoscopic applications to mono applications for this type of training [18]. Luuresma conduced a study in which participants were exposed to either a monoscopic or a stereoscopic condition for an anatomy learning module. Participants with the stereoscopic condition performed better in answering anatomical questions [18].

While overall there seems to be a trend toward the benefits on performance of stereoscopic applications [34], many studies conducted to this day have yielded mixed results.

All the studies discussed in this chapter have been important in paving the way to evaluate the differences between monoscopic and stereoscopic representations. Nevertheless, one of the limitations, of most of these studies, is the complexity of tasks completed by participants. This complexity almost certainly added to the mixed results received. It is important to simplify the tasks to really understand the positive or negative effect stereopsis might have in medical contexts. For example, surgery tasks involve the recognition of shapes, the relative judgment of positions of tasks and objects, and navigational tasks. But these tasks are measured together when assessing the performance of surgical stereoscopic applications. The focus of this work is to address basic research questions around visual spatial tasks, by themselves, in medical contexts.

2.4 Research Questions

Based on the literature review the following research issues were identified:

Determine the conditions where stereoscopic displays offer an advantage over monoscopic displays for visual spatial tasks

A lot of the studies conducted have presented mixed results when comparing stereoscopic and monoscopic displays. There is a need to take a step back, and to isolate the different depth cues that stereoscopic and monoscopic displays offer to a visual spatial medical task.

Determine if individual differences play a role in visual spatial medical tasks

There is a push in the medical field to understand how individual skills may predict future success, whether for a general practitioner or surgeon. It is not clear if this success is due to inherit individual skills, practice and learning, or technology and training. It is important to identify if individual differences play a role in visual spatial medical tasks.

CHAPTER 3. METHODOLOGY

Three studies were conducted to address the research questions posed in the previous chapter. These studies measured performance of a visual spatial computer task using medical imaging data. Participants assessed the relative positions of three different objects located inside a patient's data. This task was designed for two main reasons: 1) it is similar to studies in literature that evaluated performance of traditional 2D displays in non-medical domains [47-52] and 2) it's a fundamental skill needed for diagnosis, treatment, and training.

The first study consisted of static views and recognition of the position of color objects. The second study consisted of static views using gray objects. The third study consisted of animated views of color objects. In all of the studies the task was the same: to select which of two objects was closest to a reference object. Table 1 shows a summary of the three studies.

Table 1. Summary of the Three Studies

Study	Tasks	Distances used	Animation	Color
Study 1	Visual spatial task	5-30mm	No	Color Objects
Study 2	Visual spatial task	5-30mm	No	Grays Objects
Study 3	Visual spatial task	5-30mm	Yes	Color Objects

3.1 Procedure

The studies followed the Institutional Review Board (IRB) guidelines: Participants were introduced to the study, given the activities to be completed during the session, and told that the study was voluntary and they could stop participation at any point. Participants were compensated in two out of the three studies.

The study consisted of the following:

- A pre-survey
- Visual spatial paper tests
- Visual spatial computer task to test performance

The pre-survey asked general demographic and background questions such age, gender, and previous experience with stereoscopic applications. This pre-survey was used to identify possible variables that could influence the results of the study. For a complete list of questions, please see Appendix.

The visual spatial paper tests used were the Mental Rotation Test [53], and the Paper Folding test [54]. These tests have been used in the past as a an indicator of accuracy of surgical tasks [30, 55]. Wanzel et al. [55] conducted visual spatial tests and surgical procedures of medical residents. Residents with higher visual spatial scores had higher performance scores in surgical procedures. Accuracy of visual spatial tests could help determine performance of surgical skills.

The Mental Rotation Test is a 20-item questionnaire that measures high-level visual spatial skills. The participant is presented with one object on the left side of the test, the reference object, and four objects on the right. Two of the four objects on the right are the same as the reference object on the left, presented at different angles as shown in Figure 11. The participant must pick which two of the four objects are the same as the reference object. The score is calculated by summing the correct answers minus a fraction of the incorrect answers [53].

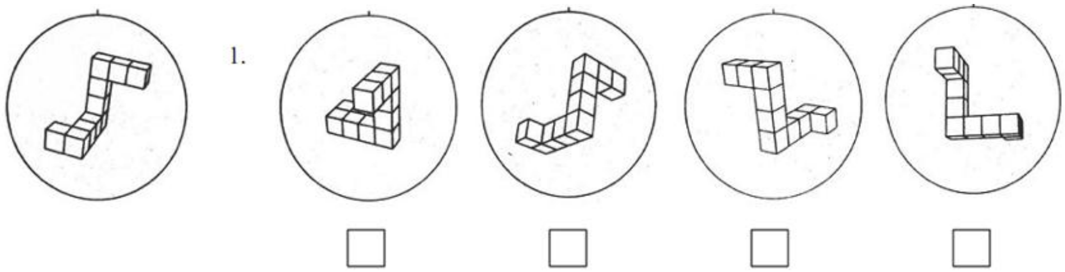


Figure 11. Standard Question of the Mental Rotation Test.

The second visual spatial test used was the Paper Folding Test. In this test the participant is asked to imagine the folding and unfolding pieces of a paper while it is being punctured. The participant is shown a figure that has been folded and punctured, and is then asked which of five figures represent the paper with holes after it has been unfolded as shown in Figure 12. As with the MRT, the score is calculated by summing the correct answers minus a fraction of the incorrect answers.

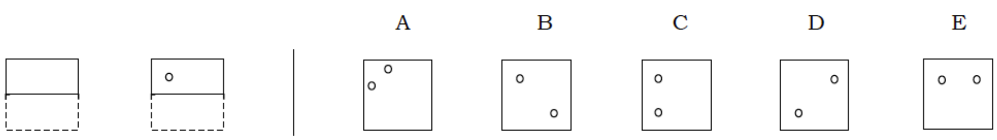


Figure 12. Example of a Paper Folding test question.

The visual spatial computer task consisted of selecting which of two cylinders were closer to a reference cylinder. Three cylinders were introduced into real patient data; the cylinders represented the basic shape of instruments that one may find in minimally invasive surgical procedures. One cylinder was set at the center of the data for every task, and was considered the reference cylinder. The other two cylinders were placed around the reference cylinder, with one of the cylinders always being closer to the reference cylinder, as shown in Figure 13. To ensure that tasks had varying levels of difficulty, the positions of the spheres and the cylinders were determined in consultation with anatomy professors.

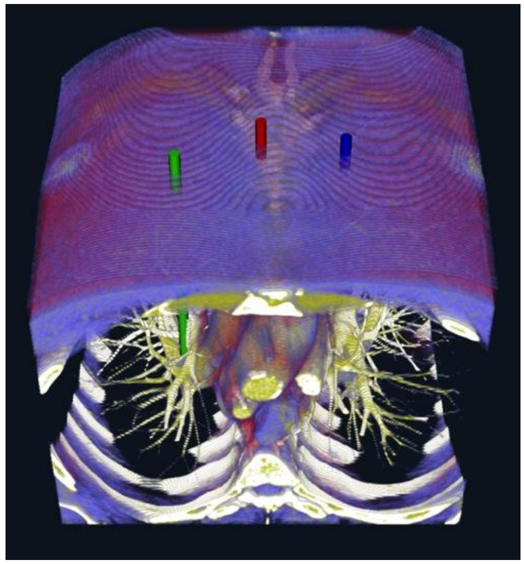


Figure 13. Example of the computer visual spatial task. A reference cylinder found at the center, and two cylinders placed around the reference cylinder.

Participants had 30 seconds to make a decision before the interface would automatically move to the next task. The 30 seconds allotted per task was selected as a consensus from literature [48]. The pilot studies conducted confirmed that 30 seconds would be enough time to make a selection.

The independent variables of the studies were the display condition: monoscopic, stereoscopic, and the distance between the cylinders. Participants only received either the monoscopic or stereoscopic condition to account for learning effects and fatigue.

Four different distances, called cases, were used. The first distance, case #1, had 30 mm between the cylinders. The second distance, case #2, had a distance of 15 mm. Cases #3 and #4 had distances of 5 mm and 2 mm, respectively.

A pilot study was used to select these distances. Initially with the first pilot study, distances between 30 mm - 15 mm were used, however these distances were too easy for the participants, with accuracy ranging from 75% to 100%. The range in distances was then increased from 30 mm – 2 mm to increase the level of difficulty.

Every answer was recorded, including the skipped answers. Data was analyzed using classic statistical techniques. When assumptions of parametric methods were met t-tests were used to compare between conditions.

3.2 Hardware Set-Up

The hardware set-up, as shown in Figure 14, was kept consistent between studies. The monitor and chair were positioned at the same height and distances for each participant.

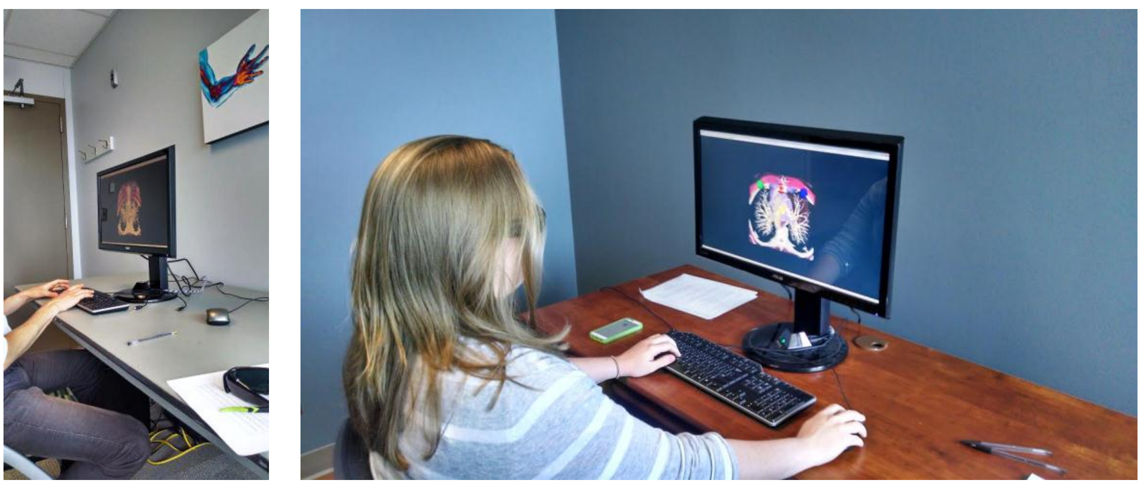


Figure 14. Set up of the user studies.

The same equipment: glasses, display, and software, was used in all three studies to minimize confounding variables.

The computer used was a Dell Precision T5500, CPU: Xeon W5580 @3.20GHz. Ram: 4 GB, nVidia Quadro FX 5800 graphics card. The monitor was a 3D Asus 23 in wide screen 3D NVIDIA ready monitor. Both the stereoscopic and monoscopic conditions used the same monitor.

The glasses used for the stereoscopic condition were NVIDIA Infrared 3D Vision wireless glasses. The participants who used the monoscopic condition also used the glasses but the stereopsis was turned off.

3.3 Software

The software used in the study, Isis, was developed in house, as a medical visualization software tool to view and manipulate digital medical images in both desktop and immersive environment [56]. For these three studies the desktop version was used.

Isis was designed to display any DICOM/PACS compatible image data. Figure 15 shows the Isis interface of abdominal patient's data. Isis allows for different types of interaction with the medical data such as: windowing, rotating, scaling, clipping, pseudo-coloring, and zooming.

Figure 15. Isis desktop software with medical data.

Windowing is the process of mapping the range of raw medical imaging values into grayscale values. Medical images usually store the data in Hounsfield Units (HU), which is a measure of tissue density relative to the density of distilled water. HU units range from -1000 HU to +2000 HU. A window center and a window width are set according to the tissue that wants to be visualized. Any HU value smaller than the window width is set to black, and any HU value bigger than the specified window width is set to white. The values within the width are set to the corresponding gray intensity. An illustration of the windowing process is shown in Figure 16, and an example of windowing for medical data is shown in Figure 17. In these studies windowing was kept constant between studies and participants were not allowed to change the settings. Bone, muscle, and fat were all present in the images.

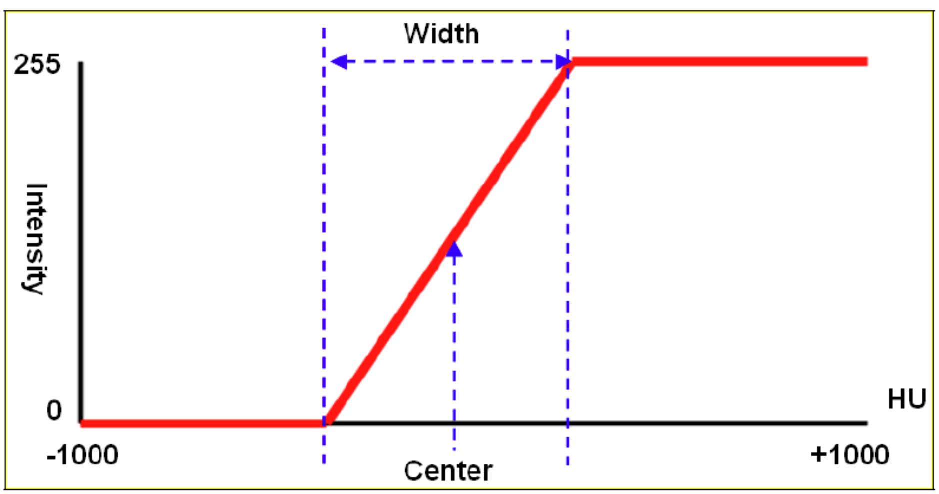
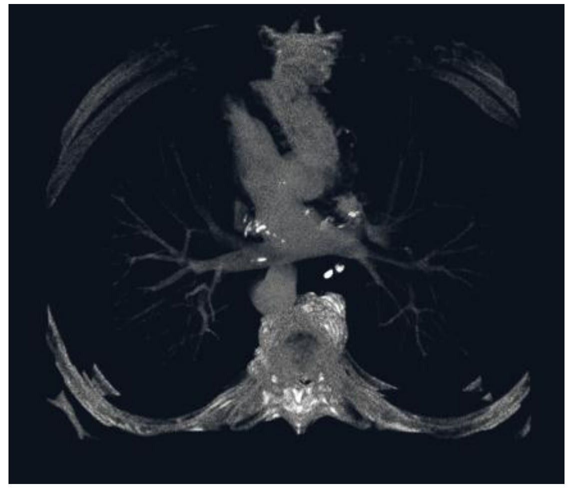


Figure 16. Illustration of windowing process for medical images.

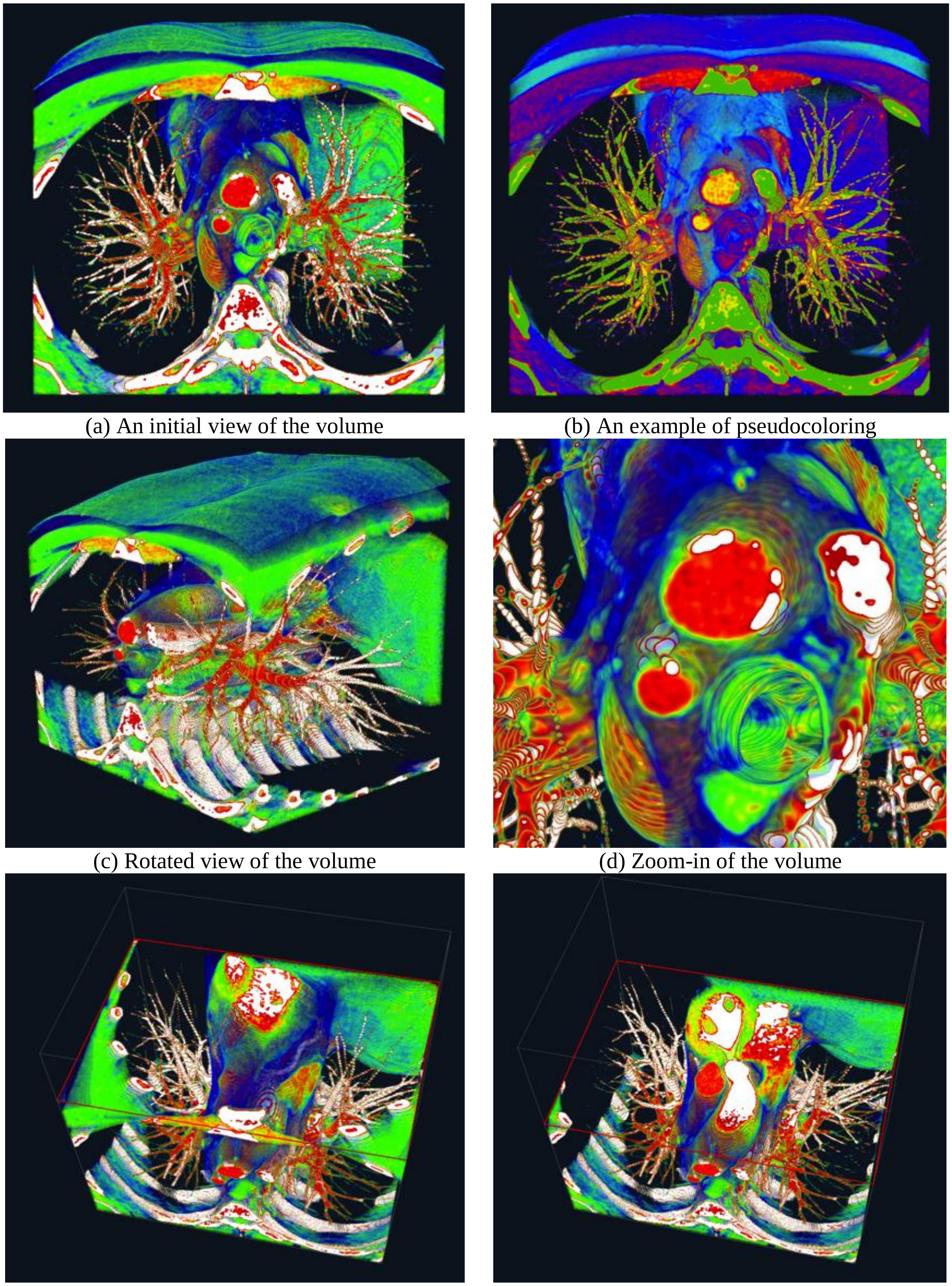
(a) Wide window width



(b) Different window width of the same data

Figure 17. Examples of different windowing settings on medical dataset.

Pseudo-coloring is the application of different coloring schemes based on pixel intensity after windowing. Different color schemes can be applied to the same volume data to highlight different regions of interest for the user, examples of the different color schemes are shown in Figure 18a-b. Rotation is the circular movement of the volume about a given point as shown in Figure 18c [57]. Zooming involves scaling the size of the volume by the origin, an example of patient’s data zoomed in is shown in Figure 18d. Clipping planes are used to create a ‘slice’ of the volume that allows for an inside view of the volume, Figure 18e-f show two examples of slicing data. These interaction techniques were kept constant during the studies to isolate any confounding variables.



(a) An initial view of the volume

(b) An example of pseudocoloring

(c) Rotated view of the volume

(d) Zoom-in of the volume

(e) Clipping plane to the top side of the volume

(f) Another view of a clipping plane cutting through the volume

Figure 18. Examples of medical visualizations interactions that can be applied to medical imaging data.

3.4 Hypothesis

The three studies tested the following hypotheses:

- H1: Adding stereopsis as a depth cue will have a positive effect on performance of a visual spatial task. This hypothesis extends from the work of previous studies on performance of stereoscopic and monoscopic applications [37].
- H2: Performance will decrease as the distance between objects being judged decreases. The stereoscopic condition will yield higher performances in the most "difficult" cases (i.e. smaller distances).
- H3: Different object orientations will have an effect on performance. Orientations at 45^{o} degree angles will result in decreased performance. The stereoscopic condition will have higher performance than the monoscopic condition at these angles.
- H4: The stereoscopic condition will have a positive yet limited influence on performance of animated visual spatial tasks. It is expected that animation affording several points of views of the task, will limit the benefits of the stereoscopic condition over static views. This hypothesis extends from previous work on performance of perspective views and animation [52].
- H5: Demographics such as age and gender will have an effect on performance of visual spatial tasks. The stereoscopic condition will have an overall positive effect across different demographic groups.

- H6: Higher performance in paper visual spatial tasks will correlate with higher performance on the computer visual spatial task. The stereoscopic condition will have an overall greater effect on lower performing individuals than the monoscopic condition. This hypothesis extends on previous work on the correlation of paper tests scores and surgical ability [55, 58].

CHAPTER 4. ASSESSMENT OF VISUAL SPATIAL SKILLS IN MEDICAL CONTEXT TASKS WHEN USING MONOSCOPIC AND STEREOSCOPIC VISUALIZATION

A paper published in SPIE Medical Imaging.

Marisol Martinez Escobar, Bethany Junke, Kenneth Hisley, David Eliot, Eliot Winer

Abstract

The dramatic rise of digital medical imaging has allowed medical personnel to see inside their patients as never before. Many software products are now available to view this data in various 2D and 3D formats. This also raises many basic research questions on spatial perception for humans viewing these images. The work presented here attempts to answer the question: How would adding the stereopsis depth cue affect relative position tasks in a medical context? By designing and conducting a study to isolate the benefits between monoscopic 3D and stereoscopic 3D displays in a relative position task, the following hypothesis was tested: stereoscopic 3D displays are beneficial over monoscopic 3D displays for relative position judgment tasks in a medical visualization setting. The results show that stereoscopic condition yielded a higher score than the monoscopic condition, but the results were not always statistically significant.

4.1 Introduction

Visualizing and understanding digital medical data can be challenging. The information generated from CT, MRI, and other medical imaging technologies can be highly detailed and complex. Translating the data into usable information is difficult, yet critical to patient's care [59]. It is estimated that the error rates in diagnosis can be up to 15% for false positives, and between 20-30% for false negatives [4]. It is crucial that medical imaging data is represented in such a way as to minimize these types of errors.

A proposed solution is stereoscopic applications. Some of the cited potential benefits of stereoscopic applications are an improvement between the foreground and the object of interest, improved image quality, and improvement in depth judgments [33]. Improving the ability to judge depth may provide an advantage in making spatial judgments in medical images. In fact the ability to see depth is so important that the Human Visual System (HVS) has redundant cues to detect it. There are several monocular and binocular depth cues that help perceive depth. Monocular depth cues can be acquired by only one eye, and these types of cues can be present in conventional displays. Monocular depth cues include perspective projection, occlusion, familiar size, object motion, and motion parallax [4, 7, 59].

Perspective projection is the projection of a 3D scene onto a flat plane and makes use of sizes, lines, and textures. As in real life, objects that are closer to the observer are projected in bigger sizes than objects that are far away. Similarly, lines that are farther apart are closer to the observer, and lines that converge are farther away from the observer. Texture patches have larger and more widely spaced elements when they are located closer to the observer. Occlusion is another monocular cue in which objects in the front block objects that are farther apart. Familiar size is another important monocular depth cue that uses a known object as a reference. This familiar object gives reference points for the sizes and depths of other objects in the scene.

Object motion, commonly obtained by rotating an object around its axis, is a monocular cue that allows the observer to see the object from different angles and construct a 3D representation. Usually the observer is static while the object is in motion.

Motion parallax is the opposite. The observer is moving while the object, or the scene remains static. The observer understands depth by the relative motion of the objects in the scene. Closer objects will move longer “screen” distances than objects which are farther away. For example, when driving a car, trees in the scenery that are closer to the car seem to move much faster than the mountains in the distance, allowing an observer to understand that the mountains are farther away. While the monocular depth cues described above are important in the understanding of depth, none of these monocular cues are as sensitive as stereopsis.

Stereopsis is a binocular depth cue that consists of processing the disparity between the unique images presented to each eye. Human eyes are slightly separated horizontally and receive a unique perspective of the world. The brain fuses the left and right images, and from the differences between the two images, the brain extracts depth information [21]. The disparity of information obtained from the two unique images is processed to the point that the HVS can differentiate between an object 1.0 m away and a second object 1.2 mm farther away [4]. Even if stereoscopic applications use the most sensitive depth cue, there wide adoption of stereoscopic applications in medicine will not occur until a clear advantage of stereoscopic applications over monoscopic applications is demonstrated.

The work presented here attempts to address this issue by studying the differences between monoscopic 3D and stereoscopic 3D displays in judging spatial relationships. How would adding the stereopsis depth cue affect relative position tasks in the medical context?

4.2 Background

Current stereoscopic medical applications can be divided into four areas: diagnosis, pre-operative planning, minimally invasive surgery, and teaching/training [4, 59].

Diagnosis aims to identify anatomical structures and abnormalities inside the human body. Currently most diagnosis applications use a mixture of 2D views and monoscopic 3D views. By making use of stereopsis to separate the background object from the object of interest, stereoscopic applications could help to reduce false positives and false negatives during diagnosis. Several stereoscopic applications have been developed for diagnostic purposes, including applications in mammography [38].

Pre-operative planning finds the optimal surgical path for a procedure based on the pre-operative images obtained from the patient. Multiple imaging and types of displays can be used simultaneously. Pre-operative planning can improve surgical precision, reduce the time of an intervention, and reduce possible complications. Similarly to diagnosis, pre-operative planning uses a mixture of 2D and monoscopic 3D views. 2D views can help to assess the distances, and angles of the surgical path, while 3D views can help to understand and label the anatomical structures [40]. Several stereoscopic applications have been developed for pre-operative planning, Reitinger [41] developed a stereoscopic liver surgery planning that allows the user to analyze the data, and conduct measurements.

MIS is another potential application for stereoscopic displays. In MIS surgery is conducted from outside the body by minimizing the size of the incisions. MIS has been proven to lower the recovery times for the patient [42]. MIS is challenging for the surgeon because visual information is restricted. The operative space cannot be seen directly, as it is only available through a camera placed through one of the incisions, called the laparoscope. The image obtained from the laparoscope is restricted to a 2D flat monitor and lacks many of the depth cues associated with open surgery [31].

Computer simulations and visualizations have been used to facilitate training. These simulations offer an alternative to cadavers, who are expensive and do not represent living tissue, and to real patients, who are not an option through part of the training. Stereoscopic applications could aid in training by providing a better spatial understanding of anatomical features. Several studies show an advantage of stereoscopic applications to mono applications for training. Luuresma [18] conduced a study in which participants were exposed to either a monoscopic or a stereoscopic training condition for an anatomy learning module. Participants with the stereoscopic condition performed better in answering anatomical questions.

While the many possible stereoscopic applications may highlight the benefits of stereoscopic displays the majority of studies conducted to this day have yielded mixed results. Kikuth [39] conducted a study for the classification of fractures using monocular 3D and stereo 3D displays and did not find any benefits or disadvantages of stereo 3D over mono 3D12. Out of the 62 images used for the identification fractures, 40 of those images had artificially created fractures by a surgeon. Kikuth commented that some of the artificial fractures were of atypical shape, which may have increased the level of difficulty of the task. Hanna [46] conducted a study in which four specialized operators conducted 60 tasks in 2D and stereoscopic displays and no differences between the conditions. This study had four experienced participants that could have been using monocular depth cues to complete the tasks. For these users stereopsis does not offer an advantage. However, this study did not show that stereopsis performed worse than the 2D condition, so while other monocular cues can be used in surgical tasks, using stereopsis as a cue does not lead to worse results.

These studies have been important in paving the way to evaluate the differences between monoscopic and stereoscopic 3D displays, but in order to assess the relevance of stereoscopic applications in a medical context more studies are necessary. The purpose of this work is to limit the amount of confounding variables found in previous studies: such as mixed displays, low number of participants, experienced participants, learning effects, and artificial data, in order to gain a true understanding of how stereoscopic applications work for a basic medical spatial task.

4.3 Methodology

The study assessed the participants' visual spatial ability by making relative judgments of objects inside a medical volume of real patient data. The study began with a pre-survey, followed by a set of visual spatial tests, and finally the set of relative position judgment visual spatial tasks to be completed. The pre-survey asked general background questions such as previous experience with stereoscopic applications. The pre-survey was used to identify possible variables that could influence the results of the study. The visual spatial tests consisted of the Mental Rotation Test [53], and the Paper Folding test [54]. These tests were performed to judge if high-level visual spatial skills would lead to better performance in relative position judgment tasks.

Three cylinders were introduced into an actual anonymized patient CT set rendered as a 3D volume. The cylinders represented the basic shape of instruments that one may find in minimally invasive surgical procedures. One cylinder (red) was set at the center of the data for every task, and was considered the reference cylinder. The other two cylinders (green and blue) were placed around the reference cylinder, with one of the cylinders always closer to the reference cylinder. The task asked the participants to select which of the two cylinders was the closest to the reference. Participants had 30 seconds to make a decision before the interface would automatically move to the next task. The 30 seconds allotted per task was selected as a consensus from literature and a pilot study confirmed this duration was sufficient [48]. For every task the answer of the participant was recorded, including the skipped answers.

The data set consisted of 354 slices of 512x512 CT images of the chest. This data set was selected because first and second year medical students, the target participants for the study, were not overly familiar with these anatomical structures and thus could not use additional cues to aid them in the evaluating the positions of the cylinders.

The independent variables of the first study were distances, and views. Distances refer to the different positions of the cylinders. Four different distances, called cases, were set up by varying the depth distance between the cylinders from the reference cylinder, this was similar to what is currently being done in literature [48, 52]. Case #1, was 30 mm between the cylinders. Case #2, had a depth distance of 15 mm, and case #3 and #4 had 5 mm and 2 mm respectively. It is expected that tasks with the furthest distances between the cylinders would be easier for the participant, thus making case #1 the easiest, and case #4 the hardest.

There were seven different views used (i.e. different volume orientations). The different views were up, center, down, center right, center left, up left, and up right as shown in Figure 19. The views up, center, and down, were repeated for every case. The views where the volume was at a rotated angle about the yaw (Figure 19b, 19c, 19e, and 19f), were only repeated in two of the four cases. For every view and at every case there were two mirrored images, one in which green was the cylinder closest to the reference, and the other one in which the blue was the cylinder closest to the reference. In total there were 40 tasks, covering the four cases of distances. The number of tasks was limited to 40 as to not fatigue participants. It was expected that some of the views would be easier for participants, such as the down view, because distortions are minimized at this angle, it is easier to make a relative distance judgment along a line parallel to a face. Other views such as the ones with the angles would be more difficult for participants, because the volume data could partially occlude the cylinders.

4.3.1 Equipment

The equipment used consisted of a Dell Precision T5500. CPU: Xeon W5580 @3.20GHz. Ram: 4 GB, nVidia Quadro FX 5800 graphics card. The monitor used was a Asus 23 inch Wide screen 3D NVIDIA ready monitor. For the stereoscopic condition NVidia Infrared 3D Vision wireless glasses were used.

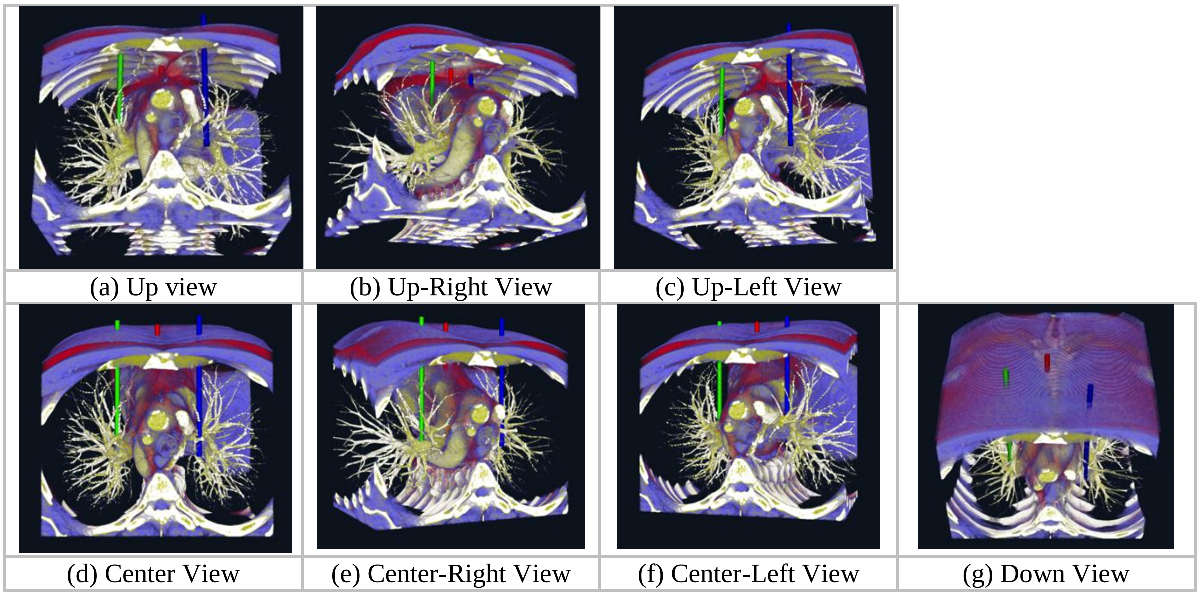


Figure 19. Different volume views used in the study.

4.3.2 Procedure

Following the Institutional Review Board (IRB) guidelines, each of the participants was introduced to the study. Participants were told the activities to be completed during the session, and that the study was voluntary and they could stop participation at any point. Participants were not compensated for their time.

There were 35 participants in the study. To avoid possible learning effects, participants were only exposed to one of the display conditions: either mono or stereo. The tasks were pseudo-randomized into three sets. Both conditions used the three sets. Eighteen of the participants went through the task in the stereoscopic condition, and 17 participants went through the monoscopic condition. One of participants from the mono condition had to be dropped from the study because their result file became corrupted.

After the collection of the results, an error was found with five of the tasks, so they were dropped, the rest of the 35 tasks were analyzed in the following section.

4.4 Results

The proportion correct (PC) score was calculated for all the cylinder tasks by dividing the number of correct answers by the total number of possible correct answers, this is shown in Figure 20. Tasks #1-10 corresponds to case #1, #11-20 corresponds to case #2, #21-30 to case #3, and #31-40 to case #4. PC ranged from approximately 12.5% to 100% depending on the task. The differences in accuracy were expected because different tasks had different levels of difficulty.

10 of the 35 tasks had 80% or higher accuracy for the mono condition, while 14 out of the 35 tasks had 80% or higher accuracy for the stereo condition. PC for two of the tasks was below 20% for the mono condition while only one of the tasks was below 20% for the stereo condition. However, the mono condition had only six tasks with PC under 40%, while stereo had eight with PC under 40%.

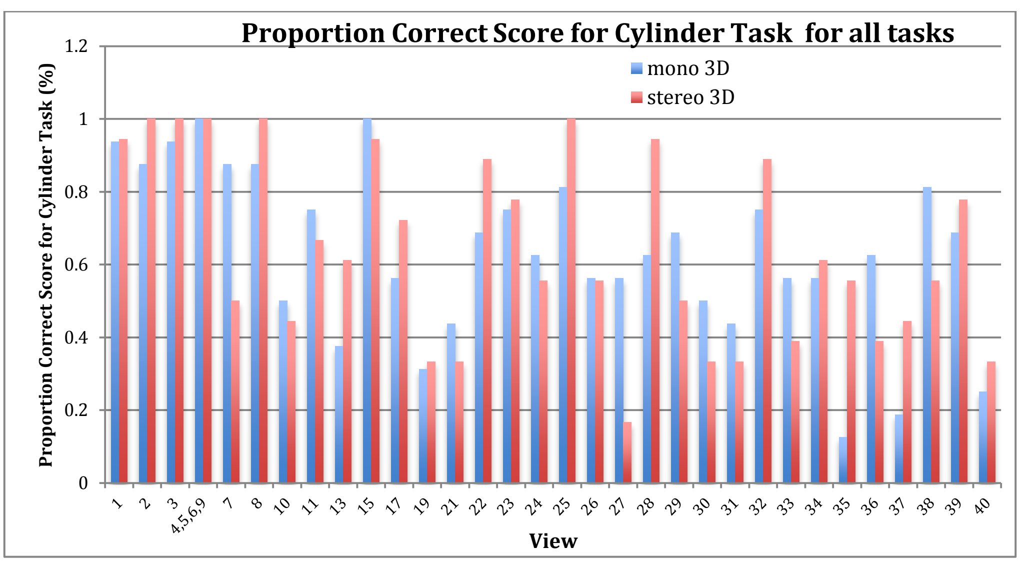


Figure 20. Proportion Correct Score for Cylinder Task across all tasks per view by mono and stereo condition.

In general, the PC score was better for the stereo condition in 18 out of the 35 views. A one-sided t-test was conducted to measure if there were statistically significant differences between mono and stereo. Five out of the 18 views were statistically significant for the stereo condition see Table 2 for the list of p-values.

Table 2. Statistically significant tasks for stereo over mono

Task Number	Volume Orientation	p-value
12	Up	0.007
16	Down	0.014
25	Down	0.029
28	Up left	0.011
35	Down	0.005

For the mono condition PC was better for 13 out of the 35 tasks, a fewer number of tasks than the stereo condition. Only two of these tasks were of statistical significance for a one-sided t-test, see Table 3, for a list of the p-values.

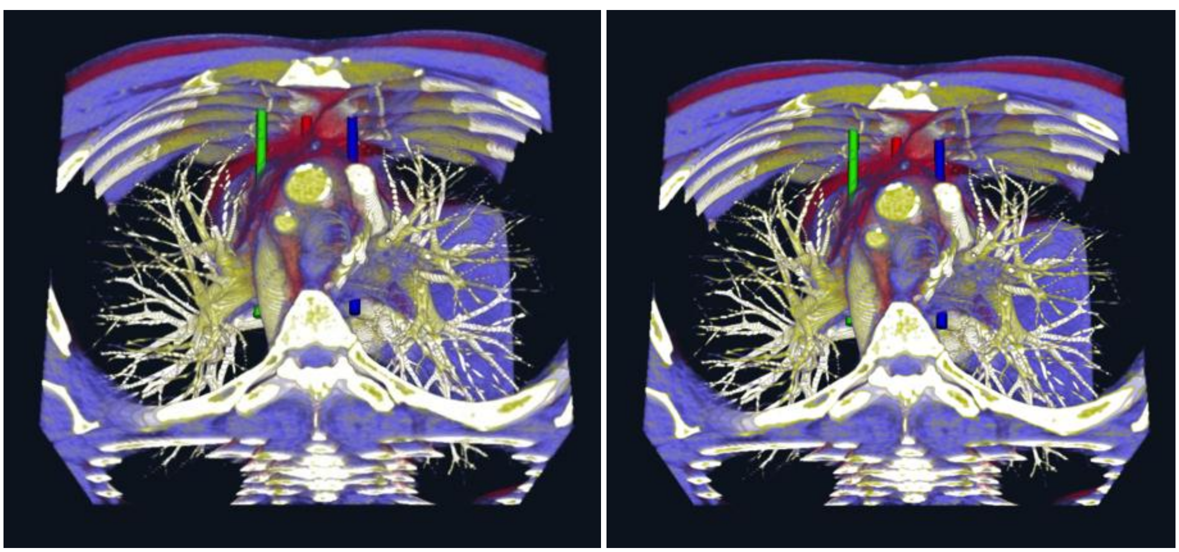
Table 3. Statistically significant tasks for mono over stereo

Task Number	Volume Orientation	p-value
7	Up Left	0.016
27	Up Left	0.007

There were some interesting patterns when the same views but different cases were grouped together. As mentioned in the methodology section, each of the up, down, and center views were repeated through the four cases. While the angled views: up left, up right, center left, and center right, were only repeated in two of the cases.

For the up view where the green cylinder, found on the left side, was closest to the reference, the mono condition was better in general, with accuracies of 93.75%, 75%, 43.75%, and 43.75% for case #1, #2, #3, #4, respectively, while the stereo condition had

PC of 94.4%, 66.67%, 33.33%, and 33.33% respectively. None of the differences were of statistical significance for a one sided t-test. For the same view with the mirror image, where the blue case was closer to the reference cylinder, the stereo condition was better with accuracies of 100%, 88.80%, 88.89%, while mono had accuracies of 87.5%, 68.75%, and 75%. View #2 was statistically significant for a one-sided t-test (p=0.035). This is interesting because the view, and the distances, were the same. The only thing that changed was the position of the closest cylinder to the reference position. Why would keeping the same view, and same distance, but different colors, yield different accuracies? Observing the images (Figure 21a) it seems that perceptually both cylinders are at the same depth, while on Figure 21b, the blue cylinder seems to be closer to the reference cylinder. There may be color perceptual implications. The results of other cases similar to this (e.g., Down View) exhibited the same behavior.



(a) Example of up view, green cylinder is closer

(b) Same up view, same distances, but blue cylinder is closer

Figure 21. Images that participant saw for the Up view task, left cylinder is the answer in a, and right is the answer for b.

The score of the paper folding test (PFT) was paired with the PC of the cylinder tasks to see if there was any correlation between high-visual spatial skills and the cylinder tests of this study. There seems to be a loose correlation between the scores of the PFT and the score of the cylinder task, the higher the score of the PFT the higher the score of the cylinder task (Figure 22). There is one outlier that scored very low on the PFT, around 20%, and scored high on the cylinder task, around 75%, but in general participants scoring higher than 30% on the PFT scored higher than 50% on the cylinder tasks. The scores of the PFT were also divided for the mono and stereo conditions. These results can be seen in Figure 23 and Figure 24. Participants who scored higher on the PFT seemed to score higher on the cylinder task. The stereo condition did not show quite the same pattern, while there still seems to be a positive correlation between the PFT and the PC of the cylinder task, the data is more evenly distributed. This could be for a number of reasons such as: 1) the PFT score is not a good indicator for the accuracy of the cylinder task, or 2) the stereo condition provided an additional cue that aided all participants. However, of particular note is that none of the results had any statistical significance. So, although a slight trend may be present it can’t be statistically verified. A similar analysis was performed for the MRT also, with the same results. No clear trends and no statistical significance.

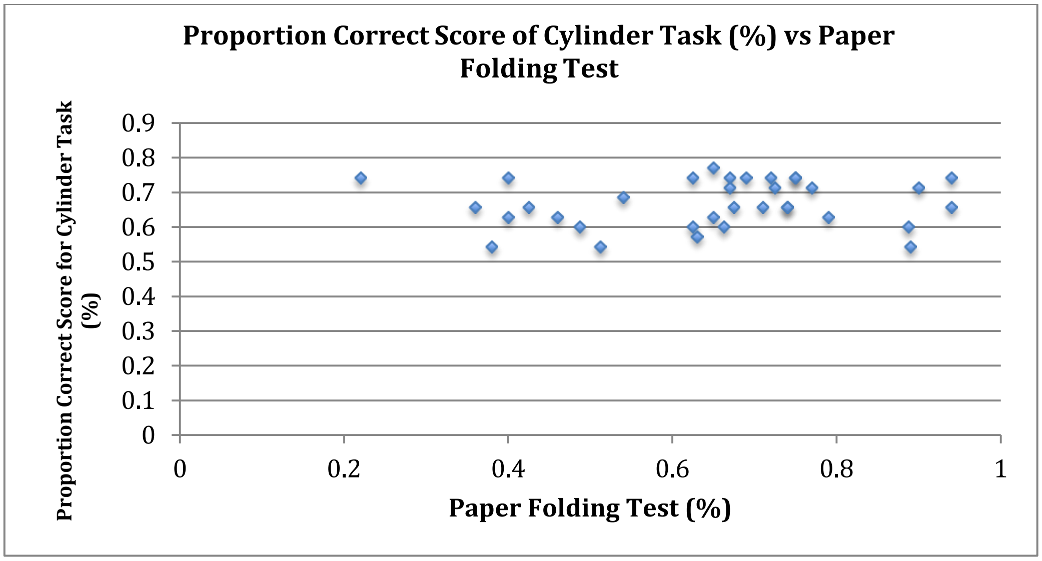


Figure 22. Proportion Correct Score for the Cylinder Task versus the Paper Folding Test Score.

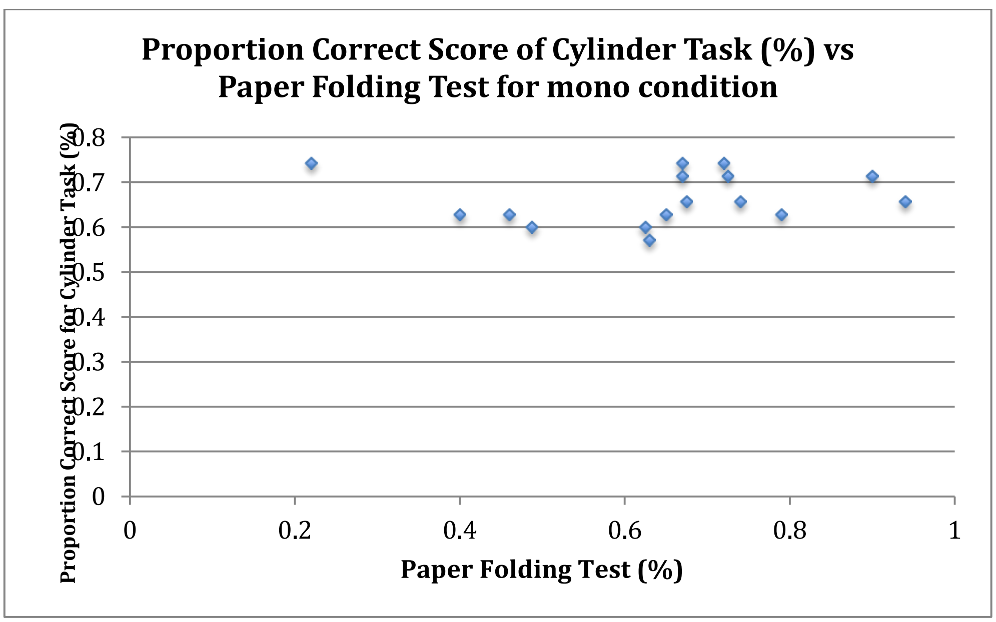


Figure 23. Proportion Correct Score for the Cylinder Task versus the Paper Folding Test Score, mono condition.

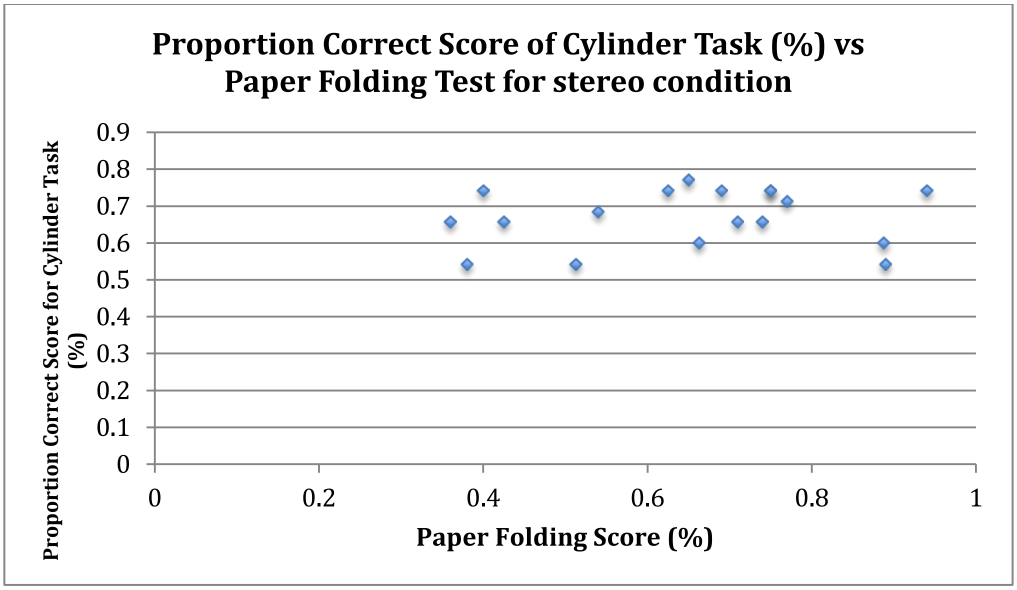


Figure 24. Proportion Correct Score for the Cylinder Task versus the Paper Folding Test Score, stereo condition.

4.5 Discussion

The stereoscopic 3D condition yielded a higher PC score than the monoscopic conditions for 18 out of the 35 tasks. The monoscopic condition had a higher PC score for 13 out of the 35 tasks. Overall, for the stereoscopic condition the PC score was 2% higher than the monoscopic condition, without any statistically significant differences for a one sided t-test.

While it was expected that the stereoscopic condition would be better than the monoscopic condition, because the stereoscopic condition makes use of the most sensitive depth cue in the Human Visual System, it was surprising to find that the monoscopic condition scored higher in some of the tasks over the stereoscopic condition.

It was also interesting to see the patterns that emerged from the data, how certain views seemed to favor the stereoscopic condition, while their mirrored images favored the monoscopic condition. This was especially apparent for the up and down view tasks.

This is interesting because this was present in all the cases, regardless of the distance. This aspect will be investigated in the remaining studies to understand what factors are being assessed in performing this task. There may be an orientation variable, but also a color perception variable. Humans perceive color differently in the brain, so having the cylinder red, green, and blue, may have affected the overall results.

The Visual spatial Tests did not show a strong correlation to the cylinder task. One of the possible problems is that the PC score of the cylinder task was not lower than 50% while the scores for the Visual spatial Test were lower than 50%. As mentioned in the background section, high-level visual spatial tasks only correlate with high-level surgical procedures, perhaps certain tasks in this study were not difficult enough to be paired with high-level visual spatial tests, and instead of using the PC score for all the tasks, the scores should have only been used for the most difficult tasks. It could also be possible that the Mental Rotation Test and the Paper Folding Test are not good at assessing relative positions tasks in a medical context.

Generally the study showed some emerging patterns in which stereoscopic displays are beneficial over monoscopic displays. The results need to be validated and investigated in future studies.

4.6 Conclusion

While stereo 3D viewing shows some promise, the results of this study are inconclusive. While overall task accuracy was slightly higher for the stereoscopic 3D views, the lack of more statistical significant measures provides evidence that more studies must be conducted to establish if and when stereoscopic 3D can aid medical personnel in everyday diagnosis and treatment decisions. This study did show that this is an issue that merits further research. 3D representations of medical data are relatively new compared to their 2D counterparts. It is important to recognize that the representations, as well as the tasks themselves, must be carefully created to maximize whichever type of visual representation is used.

4.7 Future Work

Future studies need to be conducted to validate this study and to further the understanding of the benefits of stereoscopic displays.

First, another study under the exact conditions will help to validate the patterns seen between the different views for the mono and stereo condition, and to understand how the colors, and relative positions of the cylinders play a role in the relative position task. Cylinders will not be colored, but labeled in a different manner. Second, motion or animation of the volume will be added to see if this aids task accuracy in either stereoscopic of monoscopic 3D environments.

CHAPTER 5. ADDITIONAL DISCUSSION OF VISUAL SPATIAL SKILLS IN MEDICAL CONTEXT TASKS WHEN USING MONOSCOPIC AND STEREOSCOPIC VISUALIZATION

In addition to the results reported in the conference paper additional analysis was performed to the data. This additional analysis was not included in the conference paper due to page limitations.

The performance average of all participants was 66.68%, the minimum performance value was 54.28% and the maximum performance value was 77.14%. There were no outliers in the data. A histogram and box-plot of the data is shown in Figure 25.

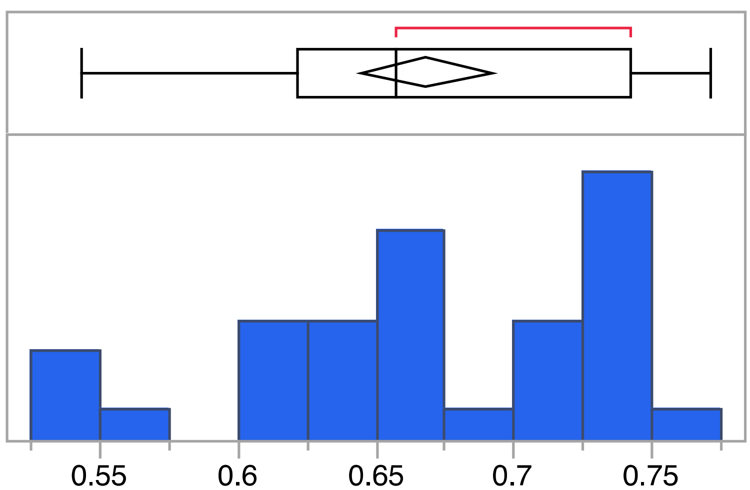
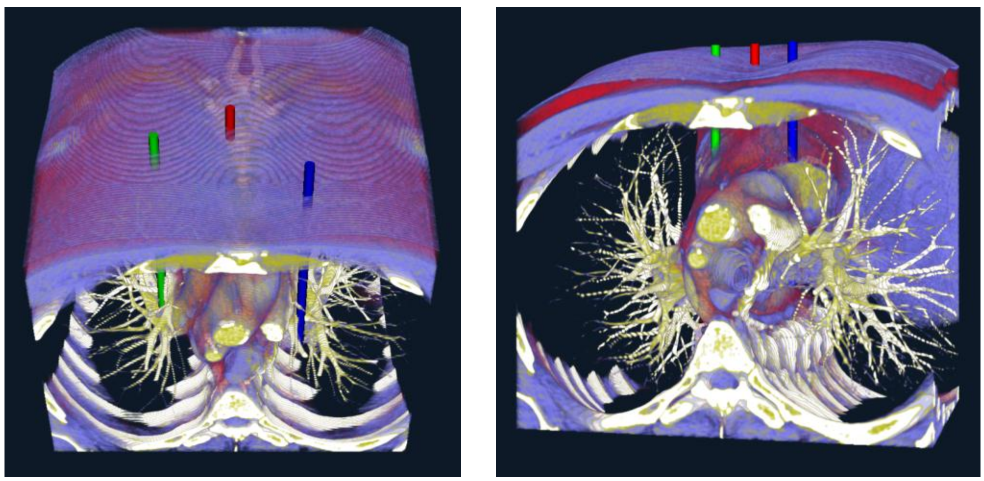


Figure 25. Distribution, box-plot and histogram for the performance with all participants.

5.1 Volume Orientation

As discussed in the methods section, there were seven volume orientations in the study. The highest average performance was the Down orientation (Figure 26a) at 75.63%. The Center Left orientation had the lowest performance at 45.10% (Figure 26b). A list of the average performance by orientation is shown in Figure 26.



a) Down Orientation b) Center left orientation

Figure 26. Highest average performance and lowest average performance volume orientations.

Table 4. Average performance by volume orientation

Orientation	Performance (%)
Center Left	44.9
Up Right	54.7
Center Right	62.1
Up Left	69.4
Center	69.7
Up	70.9
Down	75.5

Out of the seven volume orientations the performance for the stereoscopic condition was higher for five of the cases, as shown in Figure 27, with none of these differences being statistically significant. For the most difficult orientation cases, with performance average below 60% (Center Left and Up Right orientations) the stereoscopic condition had better performance. For the Center Left orientation the stereoscopic condition had an average of 48.15% and the monoscopic condition had a performance average of 41.66%, this difference was not statistically significant (p=0.111). For the Up Right orientation the performance average for the stereoscopic application was 57.41% and for the performance average for the monoscopic condition it was 52.08%, this difference was also not statistically significant (p=0.7029).

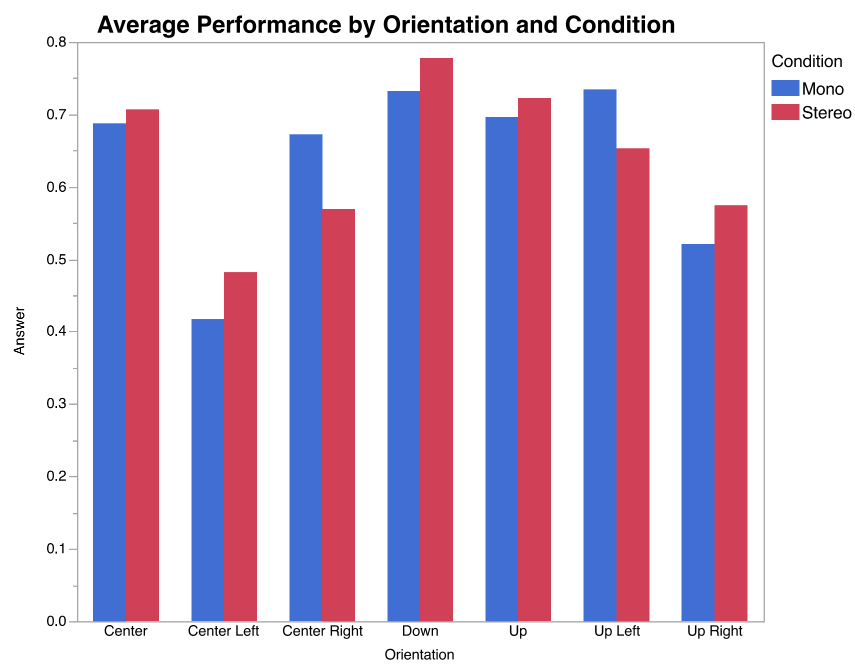


Figure 27. Average performance by volume orientation and display condition.

5.2 Demographics

Out of the 34 participants 18 were male, 15 were female, and the demographic data for one participant was not properly recorded. The average performance of male participants was 66.19% and for females it was 67.23% this difference was not statistically significant (p=0.70).

There were two age groups, participants from 18-24, and participants from 25-40. The group 18-24 had an average performance of 67.62% and the group of 25-40 had an average performance of 66.36%. This difference was not statistically significant (p=0.6604).

There were three levels of anatomy experience with the groups: participants who didn't take any anatomy classes, participants that have taken an anatomy course, and participants that have taken two or more anatomy classes as shown in Figure 28. Participants with no anatomy experience had the lowest performance, however the difference was not statistically significant (p=0.2508).

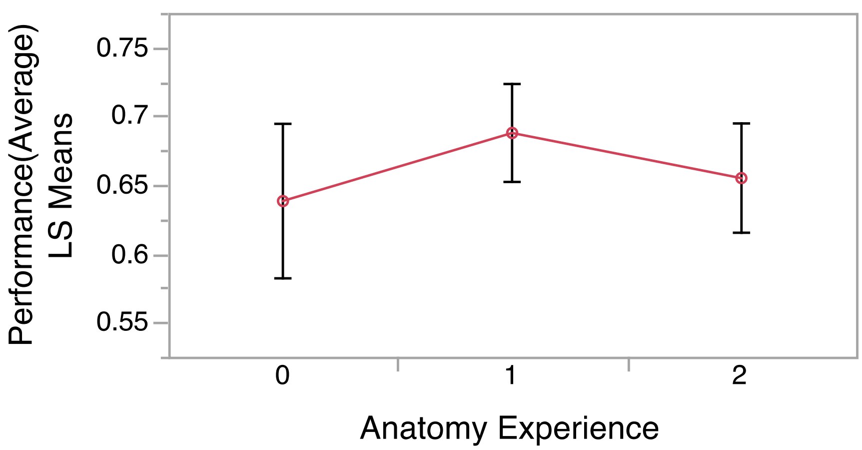


Figure 28. Average performance by anatomy experience.

5.3 Random Set

To test if the answers from the participants could have been done by chance, since participants had a 50/50 chance of selecting the right answer for each task, a random number list of answers was produced. The average between the random number generated list and the actual performance of each participant list was compared. The random list generated numbers that averaged 50.30% accuracy (as expected) while the user study average was 68.39%, with the difference being statistically significant ($p=0.0012$). There is a very strong presumption that the performance of participants was not simply by chance.

5.4 Discussion of Study 1

5.4.1 H1: Adding stereopsis as a depth cue will have a positive effect in performance of a visual spatial task.

This hypothesis was partially supported in the first study. In general the stereoscopic condition had a higher performance than the monoscopic condition. Out of these tasks, 18 of the 35 had a higher performance for the stereoscopic condition with five of the tasks being statistically significant. Of the 14 tasks with higher performance for the monoscopic condition two were statistically significant.

5.4.2 H2: Performance will decrease as the distance between objects being judged decreases. The stereoscopic condition will yield higher performances in the most “difficult” cases (i.e. smaller distances).

The stereoscopic condition had a positive benefit in performance for two out of the four cases, but it was not statistically significant. This hypothesis was not confirmed with this first study.

Performance was statistically different depending on the volume orientation. The down orientation had the highest performance. The down orientation affords a user with an almost top down view, facilitating the judgment of the positions to almost a 2D problem. Angled orientations such as center left, and up left had the lowest average performance. This is due to the perspective artifacts introduced with angles, making the assessment of positions more difficult and truly a 3D problem.

5.4.3 H3: Different object orientations will have an effect on performance. Orientations at 45^o degree angles will result in decreased performance. The stereoscopic condition will have higher performance than the monoscopic condition at these angles.

The hypothesis was confirmed as objects that were at a 45^o angle had lower performance of the visual spatial tasks. Out of the seven orientations, five of them had higher performance values with the stereoscopic condition there was not a statistically significant difference.

5.4.4 H6: Higher performance in paper visual spatial tasks will correlate with higher performance on the computer visual spatial task. The stereoscopic condition will have an overall greater effect on lower performing individuals than the monoscopic condition.

For hypothesis H6, no statistically significant correlation was found between performance on the paper tests, and the performance on the visual spatial tasks. This is different from what was found in literature. There was no pattern seen from the data, and the tests did not show correlation between the tests and the visual spatial tasks [55, 60].

5.5 Conclusions

The first user study provided a foundation to understand a basic question: how did the stereoscopic condition affect a relative position judgment task?

The results of this study show some emerging patterns in which the stereoscopic display condition yielded higher performance for most of the tasks. Additionally, the stereoscopic condition had higher performance for the most difficult cases but did not yield higher results under every case and condition.

The results of this serve as a starting point in the Human Factors of stereoscopic applications, and further research questions arise: Under which circumstances do the stereoscopic condition offer the most benefit over the monoscopic condition?

The questions left open from this initial study are addressed in the subsequent studies presented in the following chapters.

CHAPTER 6. EVALUATION OF MONOSCOPIC AND STEREOSCOPIC DISPLAYS FOR VISUAL SPATIAL TASKS IN MEDICAL CONTEXTS

A paper submitted to the Computers in Biology and Medicine Journal

Marisol Martinez Escobar, Bethany Junke, Joseph Holub, Kenneth Hisley,

David Eliot, and Eliot Winer

Abstract

In the medical field, digital images are present in diagnosis, pre-operative planning, minimally invasive surgery, instruction, and training. The use of medical digital imaging has afforded new ways to interact with a patient, such as seeing fine details inside a body. This increased usage also raises many basic research questions on human perception and performance when utilizing these images. The work presented here attempts to answer the question: How would adding the stereopsis depth cue affect relative position tasks in a medical context compared to a monoscopic view? By designing and conducting a study to isolate the benefits between monoscopic 3D and stereoscopic 3D displays in a relative position task, the following hypothesis was tested: stereoscopic 3D displays are beneficial over monoscopic 3D displays for relative position judgment tasks in a medical visualization setting. 44 medical students completed a series of relative position judgments tasks. The results show that stereoscopic condition yielded a higher score than the monoscopic condition with regard to the hypothesis.

6.1 Introduction

Visualizing and understanding digital medical data is a complex yet critical process involved in diagnosis, pre-operative planning, minimally invasive surgery, instruction and training in the medical field. The information generated from CT, MRI, and other medical imaging technologies can be high in volume, highly detailed and complex. Translating this data into usable information is difficult, yet of the uttermost importance to patients' care [59]. At 28.6%, diagnostic errors are the leading type of medical paid claim in the United States. Diagnostic errors leading to death or permanent injury range from 80,000-160,000 patients annually [61]. Diagnostic errors occur because: 1) physician cognitive bias, for example when a physician focuses in a single diagnosis without considering all the diagnostic possibilities and 2) Limited information, medical imaging data is not always available in diagnosis [62]. It is crucial that medical imaging data is represented and used in such a way as to minimize diagnostic errors.

Visualization tools and applications are being developed to minimize the complexity of medical data. One of the proposed solutions is the use of both monocular and binocular cues for depth perception, as in the case of stereoscopic applications. Depth perception is used to determine the shapes of objects in the world, and the spatial position of these objects [34]. The ability to see depth is so important that the Human Visual System (HVS) has redundant cues to detect it. Monocular depth cues can be acquired by only one eye, and these types of cues can be present in conventional displays. Monocular depth cues include perspective projection, occlusion, familiar size, object motion, and motion parallax [4, 7, 21]. If monocular depth cues are degraded or absent in traditional displays then depth perception and performance may also degrade. Stereoscopic technology can provide more robust, natural, and accurate depth perception cues in image representations.

Stereopsis is a binocular depth cue that consists of processing the disparity between the unique images presented to each eye. Human eyes are separated slightly horizontally and receive a unique perspective of the world. The brain fuses the left and right images, and from the differences between the two images, the brain extracts depth information [21]. The disparity of information obtained from the two unique images is processed to the point that the HVS can differentiate between an object 1.0 m away and a second object 1.2 mm away [4].

Even if stereoscopic applications use the most sensitive depth cue, and theoretically offer advantages over monoscopic monitors when displaying depth information to the HVS, the adoption of stereoscopic applications in medicine will not occur until a clear advantage of stereoscopic applications over monoscopic applications is demonstrated.

The work presented here attempts to study the how the stereoscopic depth cue affect a visual spatial task over a traditional monoscopic 3D display.

6.2 Background

Even though stereopsis applications offer the most robust and complete set of depth cues they are not widely adopted in the medical field. There are three main aspects contributing to barriers of adopting stereoscopic displays: 1) negative physiological displays associated with stereoscopic displays, 2) additional costs in hardware, and 3) the unclear understanding of the advantages of stereoscopic displays over monoscopic displays.

The first contributing factor against the adoption of stereoscopic displays is the association of stereoscopic displays and physiological effects. It is estimated that around two to three million persons in the US cannot see stereopsis and when presented with stereoscopic applications they experience negative reactions, such as headaches, blurred vision, and dizziness when using stereoscopic displays [20, 59].

The second contributing factor to the adoption of stereoscopic displays is cost. Stereoscopic systems require additional equipment such as glasses, special displays, and even specialized tracking systems. In some cases the computing requirements are more demanding than regular displays [4]. However, processing power has increased, and simultaneously decreased in cost, such that a stereoscopic display can be acquired for little more than a standard desktop computer.

The third contributing factor slowing the adoption of stereoscopic displays is whether they have clear advantages over monosocopic displays. Many of the studies that compare monoscopic and stereoscopic applications have yielded mixed results[59]. It is important to further the study of stereoscopic applications with basic research questions to understand their advantages and limitations.

6.2.1 Stereoscopic Applications in Medicine

Many stereoscopic applications have been developed to address current short-comings of traditional display technologies. Current stereoscopic medical applications can be divided into four areas: diagnosis, pre-operative planning, minimally invasive surgery, and teaching/training

Diagnosis aims to identify anatomical structures and abnormalities inside the human body. Currently most diagnosis applications use a mixture of 2D views and monoscopic 3D views. By making use of stereopsis to separate the background object from the object of interest, stereoscopic applications could help to reduce false positives and false negatives during diagnosis. Several stereoscopic applications have been developed for diagnostic purposes, including applications in mammography [63-65].

Pre-operative planning finds the optimal surgical path for a procedure based on pre-operative images obtained from the patient. Pre-operative planning can improve surgical precision, reduce the time of an intervention, and reduce possible complications. Similarly to diagnosis, pre-operative planning uses a mixture of 2D and monoscopic 3D views. The 2D views help to assess the distances and angles of the surgical path, while 3D views help to understand and label the anatomical structures. Several stereoscopic [41, 66] applications have been developed for pre-operative planning. For example, Reitinger [41] developed a stereoscopic liver surgery planning that allows the user to analyze the data, and conduct measurements utilizing pre-operative planning.

Minimally invasive surgery (MIS) is another potential application for stereoscopic displays. MIS surgery is conducted from outside the body by minimizing the size of the incisions. MIS has been proven to lower the recovery times for the patient [42]. MIS is challenging for the surgeon because visual information is restricted. The operative space cannot be seen directly, as it is only available through a camera placed through one of the incisions, called the laparoscope. The image obtained from the laparoscope is restricted to a 2D flat monitor and lacks many of the depth cues associated with open surgery.

Computer simulations and visualizations have often been used to facilitate training. These simulations offer an alternative to cadavers, which are expensive and do not represent living tissue. Real patients are often not an option while training. Stereoscopic applications could aid in training by providing a better spatial understanding of anatomical features. Several studies show an advantage of stereoscopic applications to monoscopic applications for training [18, 67]. Luuresma [18] conduced a study in which participants were exposed to either a monoscopic or a stereoscopic training condition for an anatomical learning module. Participants with the stereoscopic condition performed better in answering anatomical questions.

While the many possible stereoscopic applications may highlight the benefits of stereoscopic displays, the majority of studies conducted to this day have yielded mixed results. Kikuth [39] conducted a study for the classification of fractures using monocular 3D and stereo 3D displays and did not find any benefits or disadvantages of stereo 3D over mono 3D. Out of the 62 images used for the identification fractures, 40 of those images had artificially created fractures by a surgeon. Kikuth commented that some of the artificial fractures were of atypical shape, which may have increased the level of difficulty of the task. Hanna conducted a study in which four specialized operators conducted 60 tasks in 2D and stereoscopic 3D displays and no differences between the conditions. This study had four experienced participants that could have been using monocular depth cues to complete the tasks. For these users stereopsis might not offer an advantage due to their vast experience. However, this study did not show that stereopsis performed worse than the 2D condition, so while other monocular cues can be used in surgical tasks, using stereopsis as a cue does not lead to worse results.

These studies have been important in paving the way to evaluate the differences between monoscopic and stereoscopic 3D displays, but in order to assess the relevance of stereoscopic applications in a medical context additional research is necessary.

McIntire et al [29] conducted a review of all the stereoscopic studies available in literature today. Out of the 162 publications and 182 experiment reviews, they found that 60% had positive stereoscopic results, 15% showed unclear benefits, and 25% showed no benefits.

While much research has been done to study stereoscopic applications, it is still unclear as to precisely when they benefit a user the most over a monoscopic application. This becomes even more murky when considering this in the medical field as many studies are not run using medical tasks or imagery. It is important to understand the limitations of stereoscopic displays. The study presented in this paper was conducted under several controlled conditions to understand how these conditions affect performance in a medical context. Distances, orientations, and colors were independent variables used to assess the performance of a visual spatial task under monoscopic and stereoscopic conditions.

Based on surveyed literature the purpose of this work is to test the following hypothesis: Adding stereoscopic as a depth cue will have an overall positive effect on a visual spatial task in a medical context. The benefits of stereoscopic displays will be complex and dependent on task difficulty, and view angle.

6.3 Methodology

This study assessed participants' visual spatial ability to make relative judgments of objects inside a medical volume of real patient data. The study began with a pre-survey, followed by a set of visual spatial tests, and finally a set of relative position judgment visual spatial tasks. The pre-survey asked general background questions about previous experience with stereoscopic applications, about visualization tools, and discussed survey materials. These questions were meant to identify possible variables that could influence the results of the study.

The visual spatial tests consisted of the Mental Rotation Test [53] and the Paper Folding test [54]. These are two widely validated tests that measure high-level visual spatial skills. It was hypothesized that high-level visual spatial skills will lead to better performance in relative position judgment tasks.

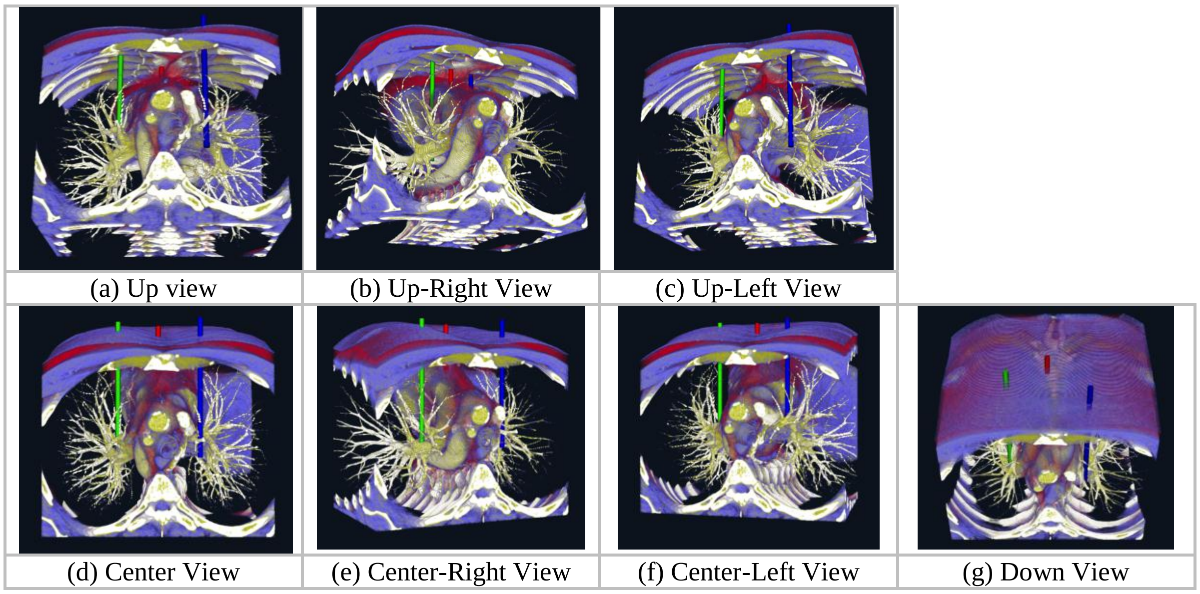
The relative position task consisted of identifying which of two cylinders was closest to a reference cylinder, similar to previous studies regarding 2D and performance [49, 54]. Sando [52] measured the difference in performance between monoscopic 2D views and monoscopic 3D views (perspective views). Sando's studies consisted of placing three objects in a white space, and asked participants to complete relative judgment tasks.

In this study, three cylinders were introduced into an actual patient CT set rendered as a 3D volume, using Sando's experimental design as a foundation. The cylinders represented the basic shape of instruments that one may find in minimally invasive surgical procedures. One cylinder was set at the center of the data for every task, and was considered the reference cylinder. The other two cylinders were placed around the reference cylinder, with one of the cylinders always closer to the reference cylinder. The task asked the participants to select which of the two cylinders was the closest to the reference. Participants had 30 seconds to make a decision before the interface would automatically move to the next task. The 30 seconds allotted per task was selected as a consensus from literature and a pilot study confirming this duration was sufficient [48]. Answers of each participant were recorded for each task, including those that were skipped.

The data set consisted of 354 slices of 512x512 CT images of the chest. This data set was selected because first and second year medical students, the target participants for the study, were not overly familiar with these anatomical structures and thus could not use additional cues to aid them in the evaluating the positions of the cylinders.

The independent variables of the study were distances, color, and views. Distances refer to the different positions of the cylinders. Four different distances, called cases, were set up by varying the depth distance between the cylinders from the reference cylinder. Case #1, was 30 mm between the cylinders. Case #2, had a depth distance of 15 mm, and cases #3 and #4 had 5 mm and 2 mm respectively. It was expected that tasks with the furthest distances between the cylinders would be easier for a participant, thus making case #1 the easiest, and case #4 the hardest.

There were seven different views used (i.e. different volume orientations). The different views were up, center, down, center right, center left, up left, and up right as shown in Figure 29. The views up, center, and down, were repeated for every case. Views where the volume was at a rotated angle (Figure 29b, 29c, 29e, and 29f), were only repeated in two of the four cases. For every view and at every case there were two mirrored images, one in which the green was the cylinder closest to the reference, and the other one in which the blue was the cylinder closest to the reference. In total there were 40 tasks, covering the four cases. It was expected that some of the views would be easier for participants, such as the down view, because distortions were minimized at this angle, and it appeared easier to make a relative distance judgment along a line parallel to a face (Figure 29g). Other views, such as those with the angles, were expected to be more difficult for participants because the volume data could partially occlude the cylinders.



(a) Up view (b) Up-Right View (c) Up-Left View

(d) Center View (e) Center-Right View (f) Center-Left View (g) Down View

Figure 29. Different volume views used in the study.

After finishing one set of 40 tasks, participants completed a second set. The differences between sets were the colors. In one of the sets all cylinders were gray (same tone and color). In the other set, the reference cylinder was red, the left cylinder was green, and the right cylinder was blue. The use and absence of color was to understand if color was playing a role in determining the distances of the cylinders. Every participant did both the grayscale, as well as the color sets.

6.3.1 Equipment

The equipment used consisted of a Dell Precision T5500. CPU: Xeon W5580 @3.20GHz. Ram: 4 GB, nVidia Quadro FX 5800 graphics card. The monitor used was a Asus 23 inch Wide screen 3D NVIDIA ready monitor. For the stereoscopic condition NVidia Infrared 3D Vision wireless glasses were used.

6.3.2 Procedure

Following the Institutional Review Board (IRB) guidelines, each of the participants was introduced to the study. Participants were told the activities to be completed during the session, and that the study was voluntary and they could stop participation at any point. Participants were not compensated for their time.

There were 46 participants in the study. The performance of two of the participants fell below 1.5Q of the minimum value, were considered outliers and dropped from the results. To avoid possible learning effects, participants were only exposed to either mono or stereo display, but not both. The tasks were pseudo-randomized into three sets used by both conditions. The grayscale and color conditions were counter balanced so half the participants saw the grayscale set first, and half the participants saw the color set first. Twenty-two of the participants went through the task in the stereoscopic condition, and 23 participants went through the monoscopic condition.

6.4 Results

The proportion correct (PC) score was calculated for all the cylinder tasks by dividing the number of correct answers by the total number of possible correct answers.

For the 40 tasks the PC of the grayscale condition was 72.2% and the PC for the color condition was 72.8%, there was no statistical difference between the color and grayscale conditions ($p=0.6679$). The data was further analyzed by mono and stereo condition, volume orientation, and difficulty level.

The PC of the stereo condition was 76.1% and 69.2% for the monoscopic condition. The difference was statistically significant ($p<0.0001$). The PC was higher for the stereoscopic condition for 26 out of the 40 tasks, 13 of them being statistically significant. The PC was higher for the monoscopic condition for 10 of the 40 tasks, but only two of them were statistically significant. The remaining four tasks had the same PC at 100% for both the stereoscopic and the monoscopic condition.

The results of the monoscopic and the stereoscopic conditions of the four cases of difficulty are shown in Figure 30. For all cases the stereoscopic condition yielded higher PC. For case #1, the easiest of the four cases the differences of PC for mono was 92.40% and stereo 95.24%. For case #2, and #3 the stereo condition was statistically significant, 11% higher for the stereoscopic condition than the monoscopic condition ($p=0.0002$). The PC for case #4 was 53.70% for the mono condition, and 57.90% for the stereoscopic condition, this difference was not statistically significant ($p=0.05$).

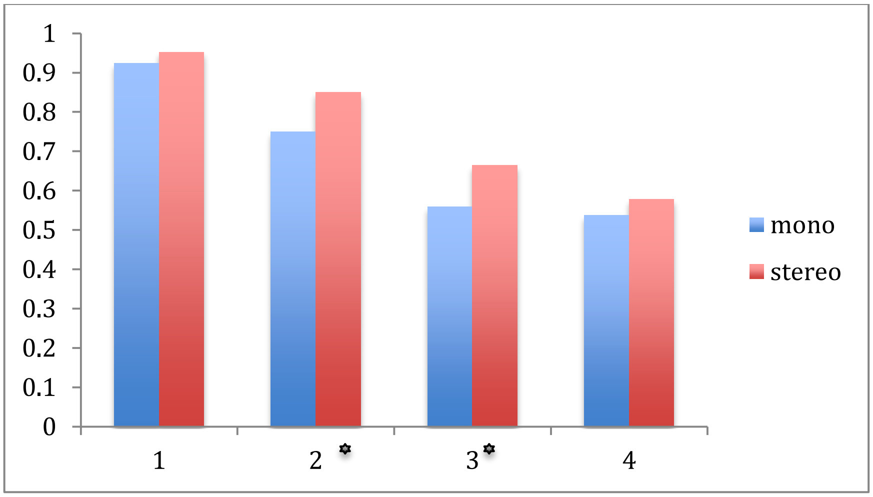


Figure 30. Results of stereoscopic and monoscopic conditions for the four different difficulty cases. *Case two and case three represent statistically significant differences between mono and stereo.

In addition to the four cases, the volume of the tasks was at seven different orientations. By analyzing the results by volume orientation the stereoscopic condition yielded higher PC for all orientations as shown in Figure 31. Four of the seven were statistically significant ($p<0.03$): the up, center, down, and up right. While the volume orientation of up left, center right, and center left orientation were not statistically significant the stereoscopic condition still yielded higher results.

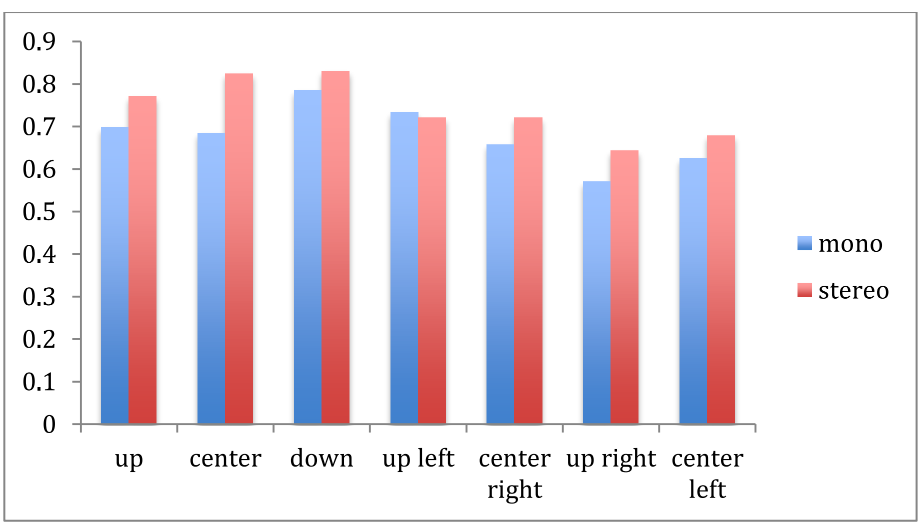


Figure 31. Results for stereoscopic and monoscopic by volume orientation

6.5 Discussion

Overall the performance difference between the color and the grayscale condition was of less than 1% and this difference was not statistically significant.

The stereoscopic condition yielded higher performance in 26 out of the 40 tasks, with 13 of these being statistically significant ($p<0.05$). The monoscopic condition yielded higher performance on 10 of the 40 tasks, with only two of them being statistically significant ($p<0.05$). Overall the stereoscopic condition yielded higher performance for the majority of the tested cases. Only in four of the tasks were the PC results the same for both the stereoscopic and the monoscopic condition. In these tasks, participants were able to score 100% in either condition.

As the distances between the cylinders decreased, increasing task difficulty, performance decreased for both the stereoscopic and the monoscopic condition. However, the stereoscopic condition yielded higher performances for the four cases, with statistically significant differences ($p<0.03$) for the second and the third case.

For case #1, the easiest case, the monoscopic condition yielded 92% performance and the stereoscopic condition yielded 95% performance, both very high. On case #4, the most difficult, the PC for the monoscopic condition was 53% and 57% for the stereoscopic condition. This difference was not statistically significant. These results suggest that for distances higher than 15mm the stereoscopic condition does not offer additional benefits than the monoscocopic condition. Between 5mm-15mm in distances, the stereoscopic condition yielded higher and statistically significant results. These may be the distances were the stereoscopic condition offers the largest benefit. Less than

5mm, the stereoscopic condition did not yield statistically significant benefit over the monoscopic condition 4.

The overall hypothesis was confirmed, stereoscopic display offers an overall benefit over monoscopic displays in relative judgment visual spatial tasks. These benefits are dependent on task difficulty. When tasks are very easy, there isn't a practical difference between the monoscopic condition and the stereoscopic condition. When tasks are of medium to hard difficulty, the stereoscopic display offers the most benefits. These findings are consistent with previously conducted research [61], where stereoscopic applications offer the most benefits with added task complexity. For the most difficult cases the stereoscopic condition was not statistically significant different than the monoscopic condition, but the stereoscopic condition did yield higher performance.

6.6 Conclusion

This paper presented a study conducted with 44 participants to assess the effect of stereoscopic displays over monoscopic displays in a visual spatial task in a medical context. Participants selected which of two cylinders were closer to a reference cylinder in patient data. Participants were 1st and 2nd year medical students.

The stereoscopic condition yielded higher performance in 20 out of the 40 tasks completed. The 40 tasks were divided into four cases, these cases used different distances between the cylinders. At distances higher than 15mm the stereoscopic condition did not yield statistically significant results than the monoscopic condition, but between 5mm-15mm the stereoscopic condition showed statistically significant higher performance. Overall the stereoscopic condition yielded higher performance across tasks, and across cases.

6.7 Future Work

Future studies to validate this work and further the understanding of the benefits of stereoscopic displays in medically sensitive tasks need to be conducted. A study to validate at what distances the stereoscopic conditions offer the most benefit for visual spatial tasks needs to be conducted. Another study that adds animation to the volume will assess how interaction plays a role in performance between monoscopic and stereoscopic conditions. These studies need to be conducted to continue the understanding of what are the benefits and limitations of using stereoscopic displays in the medical field.

CHAPTER 7. ADDITIONAL DISCUSSION OF EVALUATION OF MONOSCOPIC AND STEREOSCOPIC DISPLAYS FOR VISUAL SPATIAL TASKS IN MEDICAL CONTEXTS

7.1 Demographics

There were 44 participants in the second study with 20 male and 24 female. The average performance of male participants was 73.13% and for females was 72.03%. This difference was not statistically significant (p=0.692).

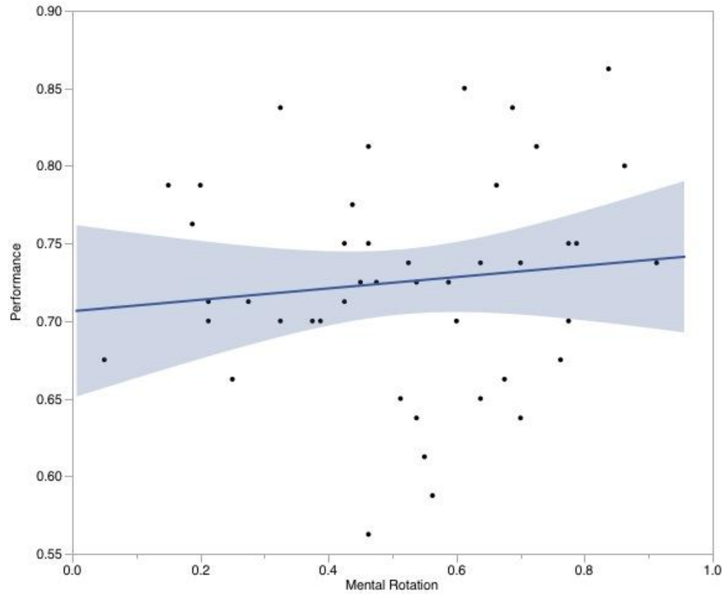
There were two age groups, participants from 18-24, and participants from 25-40. The group 18-24 had an average performance of 71.41% and the group of 25-40 had an average performance of 73.16%. This difference was not statistically significant (p=0.7762).

Analyzing the data by age group and display condition, the group under 25 had a performance average of 68.25% for the monoscopic condition and for the stereoscopic condition the performance average was 76.67%, this difference was statistically significant (p-0.0305).

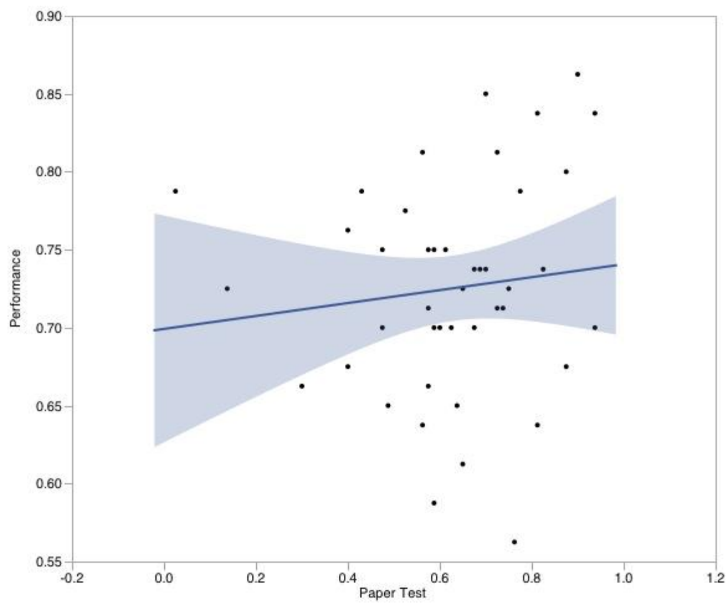
For the 25 and older age group, the performance average for the monoscopic condition was 70% and for the stereoscopic condition the average performance was 75.59%, this difference was statistically significant (p-0.0044).

7.2 Paper Tests

The score of the Paper Folding Test (PFT) test and the Mental Rotation Test (MRT) test for each participant was matched to the performance of the visual spatial tasks. There was no correlation between the paper test and the paper tasks, as shown in Figure 32.



a) Mental Rotation vs Performance



b) Paper Test vs Performance

Figure 32. Relationship between the paper tests and the visual spatial task.

7.3 Discussion of Study 2

7.3.1 H5: Demographics such as age and gender will have an effect on performance of visual-spatial tasks. The stereoscopic condition will have an overall positive effect across different demographic groups.

This hypothesis was partially confirmed on this study, there was an interaction between the performance and display condition. The performance average was higher and statistically significant for the stereoscopic condition than for the monoscopic condition for both age groups.

7.3.2 H6: Higher performance in paper visual spatial tasks will correlate with higher performance on the computer visual spatial task. The stereoscopic condition will have an overall greater effect on lower performing individuals than the monoscopic condition..

This hypothesis was not supported in this study, there was no statistically significant relationship between the score of the paper tests and the average performance of the visual spatial tasks.

CHAPTER 8. ASSESSING PERFORMANCE OF VISUAL SPATIAL TASKS BETWEEN ANIMATED MONOSCOPIC AND STEREOSCOPIC APPLICATIONS IN A MEDICAL CONTEXT

A paper submitted to the SPIE Journal of Medical Imaging

Marisol Martinez Escobar, Joseph Holub, Stacy MacAllister, Eliot Winer

Abstract

Stereopsis may be useful in medical applications as an additional depth cue to aid diagnosis, pre-operative planning, minimally invasive surgery, and training. However, previous studies show positive, mixed, and negative results when adding this depth cue. These studies did not use stereopsis in a medical context as most previous work was performed with basic manipulation or assessment type tasks of primitive objects. Thus, basic question around the benefit from stereopsis in medical applications remain. A user study to assess the performance of a visual spatial task in an animated volume of medical data between stereoscopic and monoscopic displays was completed to address this basic question. Thirty-one medical students participated in this study. The results show an overall benefit of the stereoscopic condition over monoscopic condition for this task.

8.1 Introduction

To understand images, shapes, and spatial arrangement of objects relative to each other, human visual perception involves the combination of several and often redundant depth cues [29]. These cues can be categorized as monocular and binocular. Monocular cues are obtained with only one eye such as found in conventional 2D displays (i.e.

televisions and monitors). Monocular depth cues include perspective projection, occlusion, familiar size, object motion, and motion parallax [33].

Binocular cues are obtained with both eyes. Stereopsis is a binocular depth cue that consists of processing slightly different image information received in each eye. Human eyes are separated horizontally which provides a unique perspective of the world. The brain fuses the differences between these images and calculates depth. Stereopsis is such a robust depth cue it allows the human visual system to distinguish an object that is 1.0 m away from another object that is 1.2 mm apart [4].

In medicine, stereoscopic applications are being developed with both binocular and monocular depth cues with the goal of improving the utilization of complex medical data.

Stereoscopic applications are used in diagnosis, pre-operative planning, minimally invasive surgery, and learning with positive, mixed, and negative results [7, 29]. MacIntire et al. [29] conducted a survey of studies of performance between stereoscopic and monoscopic applications. Overall 70% of the studies showed a benefit of using stereoscopic cues while 23% showed no benefit or negative benefit for stereoscopic applications. The demonstrated benefit of stereopsis was dependent on type of task and domain of the study as some studies were done with medical type tasks and some were not. For example the survey showed there were 12 total studies that studied the performance of navigational tasks with seven of these showing no or negative benefit of stereopsis. However, none of these studies were conducted in a medical context. In contrast there were 26 studies on finding and identifying objects with 17 of these demonstrating a benefit with stereopsis.

The vast majority of studies on medical applications center around the manipulation of objects, with a lack of work investigating relative judgment of position and distances [29]. Relative judgment of position and distances of objects is an important skill in medicine to understand anatomy and how organs and tissues are positioned inside the body. This is a skill needed for diagnosis, planning, surgery, and overall treatment of patients.

Given that medical errors can lead to death or permanent injury at a rate of 200,000 lives per year in the US alone [8], a clear understanding on the benefits of stereoscopic applications for visual spatial medical tasks is important. A basic research question remains on the performance of stereoscopic applications for visual spatial medical tasks.

The work presented here studies how stereoscopic depth cues affect performance on a visual spatial task in an animated volume. Previous work has assessed the effect of stereoscopic condition over monoscopic conditions using static views [68]. The current work dramatically changes the user interaction by allowing the medical imaging to be viewed from multiple angles through animation.

8.2 Background

Visual spatial ability is the capacity to understand and remember the relationships between objects. It is a skill necessary in medicine to understand the location of organs relative to each other, the shape of these organs, and how organs and tissues are connected [58].

In diagnosis, visual spatial skills are used to makes sense of a patient’s data. This entails investigating what may be problem areas, whether organs and tissues are in the right locations, and whether they are allowable sizes. In pre-operative planning, visual spatial skills are used to plan the course of action to determine what could provide difficulties in terms of sizes and positions. In minimally invasive surgery, visual spatial skills are needed to navigate, to understand relationships of organs, and to be able to move tools successfully inside the body. In training visual spatial skills are needed to understand and learn anatomy [58].

8.2.1 Stereoscopic Applications in Medicine

The stereoscopic applications in medicine developed today make use of visual spatial tasks. These applications can be divided into four areas: diagnosis, pre-operative planning, minimally invasive surgery, and training [4, 64, 65].

The goal of stereoscopic diagnosis applications is to help in identifying a patient's issues. In using stereoscopic depth cues to separate and differentiate the background objects from an object of interest, stereoscopic applications could help reduce false positives and false negatives during diagnosis. Several stereoscopic applications have been developed for diagnostic purposes [38, 64, 65]. D'Orsi et al. [65] studied the effects of stereoscopic applications in mammography. Their results show that stereopsis significantly improved diagnostic accuracy. Kikuth conducted a study for the classification of fractures using monocular 3D and stereo 3D displays and did not find any benefits of stereopsis over the monoscopic condition [39].

Stereoscopic applications in pre-operative planning help in finding the optimal surgical path before a surgery based on a patient's pre-operative images. Stereoscopic applications can improve surgical precision, reduce the time of an intervention, and reduce possible complications. Several stereoscopic applications have been developed for pre-operative planning [41, 64, 66]. Reitinger et al. [41] developed a stereoscopic application that allowed a user to analyze pre-operative data and conduct measurements of the positions between organs. Reitinger did not conduct a full evaluation study on the application, but an expert user provided positive feedback.

Minimally invasive surgery (MIS) requires high skill levels. MIS involves small movements over a narrow field of view. Visual spatial skills are needed in MIS to assess the positions of objects, and the relationship between different organs and tissues. It is difficult to judge depth over this narrow field of view. Stereopsis could provide an additional depth cue over the limited visual information in MIS. Taffinder et al. [69] conducted a study of MIS and found that a stereoscopic condition improved performance over conventional displays. On the other hand, Hanna et al. [46] conducted a study to assess performance of stereoscopic and monoscopic displays during surgery and found no difference between stereoscopic and monoscopic conditions.

Using stereoscopic applications in training and education helps students understand spatial relationships, and shapes inside the body. Several studies show an advantage of stereoscopic applications to monoscopic applications for training [18, 67]. Luuresma et al. [18] conducted a study were participants completed an anatomical training module in a stereoscopic or monoscopic condition. Participants in the stereoscopic condition had a higher performance.

Many of the stereoscopic applications in medicine have yielded positive results, however there are some studies that have yielded mixed results, perhaps because most of the current stereoscopic applications assess the overall performance in complex tasks and situations. These complex situations are actually composed of several medical skills such as navigation, relative judgment of positions, identification of objects, and object manipulation. As literature describes [34], the benefits of stereoscopic applications are dependent on the task, and complexity of the task. It is necessary to separate these complex situations into their base components to begin to understand the benefits and drawbacks of using stereopsis.

In the surveyed literature, there is little or hard to interpret data regarding the effect of stereopsis on visual spatial tasks in a medical context. This serves as the foundation of the work presented in this paper. Based on this, the following hypotheses were developed:

H1: The stereoscopic condition has a positive yet limited influence on performance of animated visual spatial tasks

- It is expected that animation affording several points of views of the task, will limit the benefits of the stereoscopic condition over static views, as studied in previous work [68]

H2: The stereoscopic condition will be most beneficial in high complexity visual spatial tasks as opposed to simpler ones

- Relatively simple visual spatial tasks (e.g., determining the relative position of simple primitives) performed in stereo will not yield better results than in a monoscopic setting. However, in tasks with increased difficulty (objects closer to organs), stereopsis will yield improved performance.

8.3 Methodology

This study assessed participants' performance in visual spatial tasks of real patient data. The study was composed of a pre-survey, a set of paper visual spatial tasks, and a set of computer visual spatial tasks.

The pre-survey asked background questions about previous stereoscopic, visualization, and medical experience. These questions were meant to identify possible variables that could act as confounding variables in the study.

The paper visual spatial tests consisted of the Mental Rotation Test [53] and the Paper Folding test [54]. These are two widely validated tests that measure high-level visual spatial skills.

The computer visual spatial task consisted of identifying which of two spheres was closest to a reference sphere during an animation. Three spheres were introduced into an actual anonymous patient CT set rendered as a 3D volume. One of the spheres was set at the center of the data set for every task, and was considered the reference sphere. The other two spheres were placed around the reference sphere, with one of the spheres always closer to the reference sphere, as shown in Figure 33.

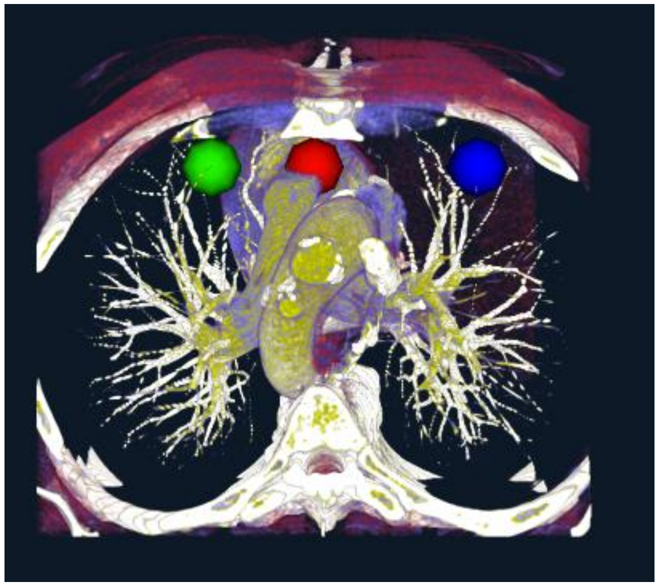


Figure 33. Example of the spheres and volume for the study.

The task asked participants to select which of the two spheres was closest to the reference sphere. In addition, each participant viewed the volume rotating along the x, y and z-axes as well as a 45^o angle axis between the primary planes. Each participant observed all rotations twice prior to selecting the sphere closest to the reference. Participants were not allowed to skip the rotation animation to ensure that every participant had the same views of the task prior to making a decision.

Following the animations, participants had 30 seconds to make a decision before the interface would automatically move to the next task. The 30 seconds allotted per task was selected as a consensus from literature and a pilot study confirming this duration was sufficient. Answers of each participant were recorded for each task, including those that were skipped.

The data set consisted of 354 slices of a 512x512 CT study. This study was selected because first and second year medical students, the target participants for the study, were not overly familiar with these anatomical structures and thus could not use additional cues to aid them in the evaluating the positions of the spheres. With limited experience in the medical field, medical students cannot use additional cues to complete the tasks, minimizing unaccounted variables in the study. Students have to rely primarily on what is presented on the screen. As the goal of the study is to assess the benefits of stereopsis in a visual spatial task as foundational research, confounding variables need to be minimized. By using students, medical experience should be similar between users minimize possible confounding variables.

The independent variables of the study were the display condition and the distances between the spheres. Three different distance ranges, called cases, were used in the study. Case #1 was 15-30 mm between the spheres. Case #2 was 5-15 mm and cases #3 was 5 mm and under. It was expected that tasks with the furthest distances between the spheres would be easier for a participant, thus making case #1 the easiest, and case #3 the hardest. Figure 34 shows a schematic of how distances were varied amongst the three spheres. In total there were 32 tasks, covering the three cases. Examples of what participants viewed are shown in Figure 35. These three cases were selected to assess at which ranges is the stereoscopic condition more beneficial than the monoscopic condition.

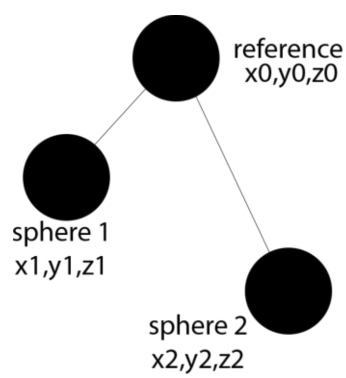


Figure 34. An example of the positions of the spheres

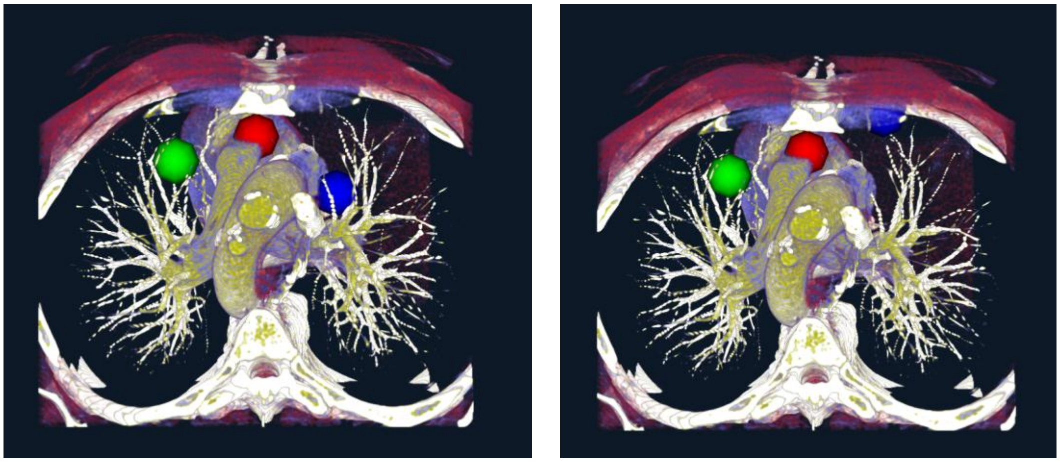


Figure 35. Examples of visual spatial tasks in the study.

8.3.1 Equipment

The equipment used consisted of a Dell Precision T5500 with an Intel Xeon W5580 CPU at 3.20GHz with 4GB of RAM and an nVidia Quadro FX 5800 graphics card. The monitor used was an Asus 23 inch Wide screen 3D NVIDIA ready monitor. For the stereoscopic condition NVidia Infrared 3D Vision wireless glasses were used. Participants used the glasses in both the stereoscopic and monoscopic conditions to minimize any confounding variables, although they had no effect in the monoscopic condition as far as affecting the visual scene presented.

8.3.2 Procedure

Following the Institutional Review Board (IRB) guidelines, each of the participants was introduced to the study. Participants were told the activities to be completed during the session, and that the study was voluntary and they could stop participation at any point. Participants were not compensated for their time.

There were 31 participants in the study. To avoid possible learning effects, participants were only exposed to one of the display conditions (i.e. mono or stereo). The tasks were pseudo-randomized into three sets. The sets had the same tasks but in different sequences to minimize effects due to fatigue and learning. Both the stereoscopic and monoscopic conditions used these three sets in equal amounts.

8.4 Results

31 participants completed the study, of which 16 in the stereoscopic condition and 15 in the monoscopic condition. Performance was measured as the % of correct answers over the total possible answers.

The average performance for all the participants was 75%, the lowest score was 56.25%, the highest score was 87.5%. There were no outliers in the study, the data of all 31 participants was used in the statistical analysis.

8.4.1 Analysis Yielding Statistical Significance

Out of the 32 tasks, the stereoscopic condition had higher performance for 14 of those tasks with four being statistically significant. These are shown in Table 5. There were no tasks for which the monoscopic condition showed statistically significant improved performance

Table 5. Statistically significant tasks for the stereoscopic condition

Task	Stereo	Mono	p-value	Case
12	81.25	40.00	0.0088	3
17	100.00	80.00	0.0314	2
27	75.00	40.00	0.0252	2
28	87.50	46.67	0.0070	2

As discussed in the methods section, there were three cases, with varying distances between the spheres. For case #2 the performance of the stereoscopic condition was 73.44% and 65% for the monoscopic condition. This was a statistically significant difference ($p=0.0321$).

The data was also analyzed by age group and gender. Participants under 25 had a performance average of 79.24% while participants over 25 averaged 71.51%. This difference was also statistically significant ($p=0.12$). Females had an average performance of 80% in the stereoscopic condition and 72.57% in the monoscopic condition, this difference was statistically significant ($p=0.074$).

The data was analyzed by display condition, age group, and difficulty case. The group of 25 and above had an average of 71.53% for the stereoscopic condition for the second case and 58.59% for the monoscopic condition, this difference was statistically significant ($p=0.022$).

8.4.2 Other results

The average performance for all participants for the stereoscopic condition was 75.35% and for the monoscopic condition was 70.34% as shown in Figure 36. This difference was not statistically significant ($p=0.1513$). There were no outliers on the stereoscopic or monoscopic conditions.

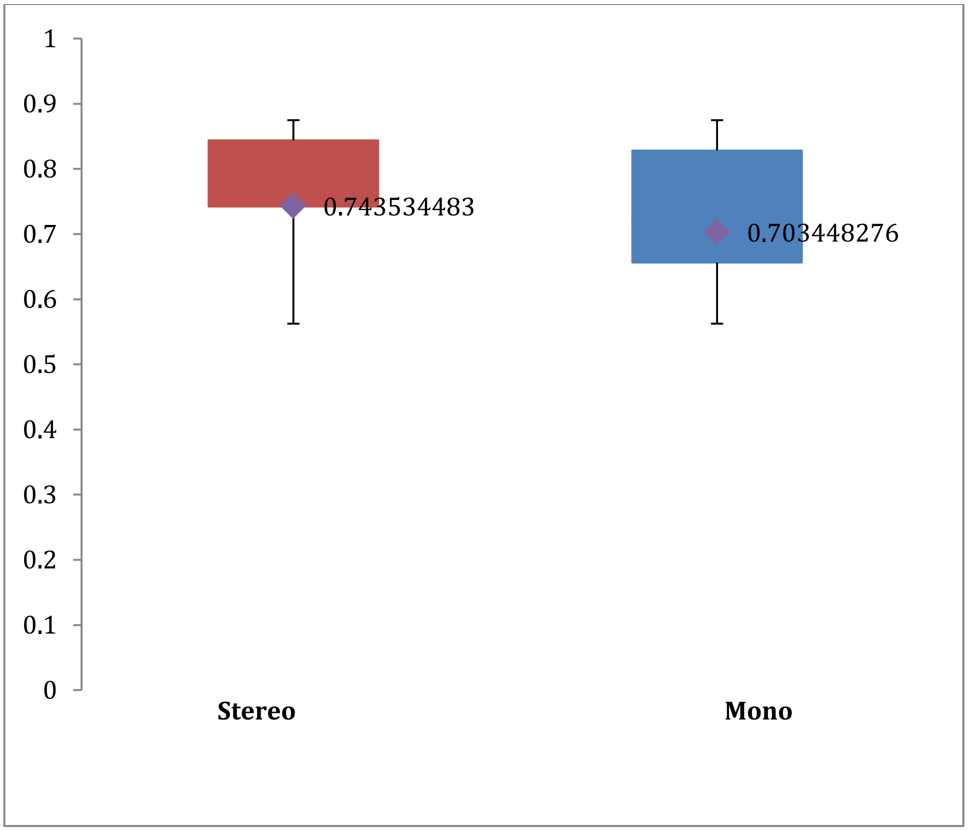


Figure 36. Box plot of stereoscopic and monoscopic conditions of all participants.

Even though the difference between the averages of stereoscopic condition and monoscopic condition were not statistically significant, there was a larger number of participants who had a higher performances on the stereoscopic condition than the monoscopic condition. Thirteen participants scored 70% or higher in the stereoscopic condition, while only nine scored 70% or higher in the monoscopic condition. On the other hand, only three participants in the stereoscopic condition had performances lower than 70%, while six participants in the monoscopic condition scored lower than 70% as shown in Figure 37.

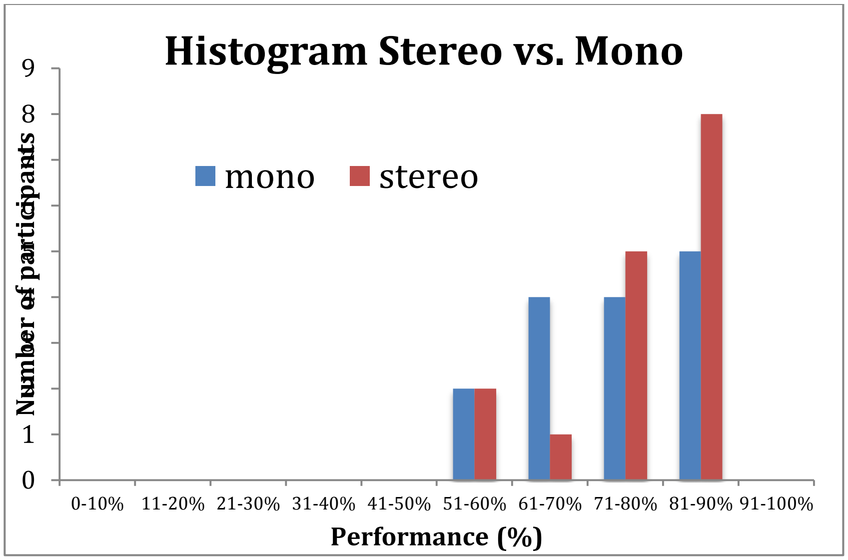


Figure 37. Histogram of performance by conditions, all participants.

8.4.3 Cases

In case #1 (easiest), the stereoscopic condition had an average of 99.22%, and the monoscopic condition had an average of 99.16%. This difference was not statistically significant (p=0.482). For case #3 the stereoscopic condition average was 51.88% while the monoscopic condition's was 53.33%. This was not a statistically significant difference (p=0,3575).

8.4.4 Demographics

Out of the participants 17 were male, and 14 were female. Females performance average was 75.22% and males was 74.82%, this difference was not statistically significant (p=0.455).

Data was analyzed by gender and display conditions. The performance average for males was 75.28% in the stereoscopic condition, and 73.96% in monoscopic condition this difference was not statistically significant (p=0.405).

Data analyzed by group and display condition yielded the over 25 age group had a performance average of 74.31% for the stereoscopic condition, and 68.36% for the monoscopic condition, this difference was not statistically significant (p=0.684). The group under 25 had a performance average of 79.91% for the stereoscopic condition, and 78.57% for the monoscocopic condition, again without statistical significance (p=0.361).

Data was analyzed by display condition, age group, and difficulty case. The group of 25 and above had an average of 98.61% for the first case in the stereoscopic condition, and 98.44% for the monoscopic condition. This was not statistically significant (p=0.467). For the third case the average of the monoscopic condition was 57.81% and for the stereoscopic condition was 55.56%, this was not statistically significant (p=0.409).

For the age group of under 25, the first case yielded 100% for both the monoscopic and the stereoscopic conditions. For the second case the stereoscopic condition yielded 75.89% and the monoscopic condition averaged 72.32% without statistical significance (p=0.27). For the third case the stereoscopic condition average was 67.86% and for the monoscopic condition it was 69.64%, this was not statistically significant (p=0.409).

8.5 Discussion

In general the stereoscopic condition yielded higher results than the monoscopic condition. Out of the three difficulty cases, case #2, displayed improved performance under the stereo condition with statistical significance. Case #2 had distances between 5-

15mm. Specifically, four out of the tasks of case #2 showed statistically significant improvement for the stereoscopic condition. None of the tasks for this case yielded statistically significant results in favor of the monoscopic condition. These distances could represent a range where stereo offers an added benefit. There are many medical procedures performed with organs, tumors, or other objects in this range.

Female participants seemed to benefit the most from the stereoscopic condition, the improvement between the stereoscopic conditions over the monoscopic condition was 8%, and it was statistically significant. This finding will need to be verified but may have implications on which gender groups may benefit the most out of stereoscopic applications, future work will need to confirm these results.

When looking at the differences between the monoscopic and the stereoscopic conditions and the age groups by cases, the group above 25 had a higher performance for the stereo condition for case #1 and case #2, with the latter being statistically significant.

This supports hypothesis 1 that the stereoscopic display condition will have a positive effect over the monoscopic condition, as in most tasks, cases, age group, and gender the stereoscopic application had an effect.

As expected while developing the study, different cases had difference performance averages. It was expected that the easiest difficulty level, case #1, would yield the highest results. And that hypothesis was confirmed in the results. Case #1 had a performance average of around 99%, case #2 had an average of 68%%, and case #3 had a performance average of 52%.

Case #2 yielded 13% higher performance average for stereoscopic as opposed to monoscopic. For the group under 25, while the stereoscopic condition yielded better results for the first and second case, it was not statistically significant. This can have implications of not only the best distances to where stereoscopic applications should be used, but also which age groups could benefit the most.

Hypothesis 2 was not completely supported. On one hand when the difficulty levels were higher, the stereoscopic condition had better performance, but not always with statistical significance. In addition, for the hardest case, the stereoscopic condition did not result in the highest performance, but this difference, again, was not statistically significant.

Synthesizing the significant and non-significant results also brings into question the effect of viewing the representation from different angles (i.e. the animation). The animation provided the participant with many angles and views that they used to assess the relative position of the spheres as well as other anatomical features. These additional angles might have provided depth cues in the assessment of visual spatial tasks lending to improved results regardless of the stereoscopic effect. This could pave the way for another study to isolate this effect.

Although the results seem to suggest overall improvement with stereopsis, the lack of clear statistical significance lends some doubt as to its overall effectiveness. The conclusion, at this point, is to add stereo if the time and effort is within reason for the application or task being performed.

8.6 Conclusion

Overall the stereoscopic condition had better performance in animated medical visual spatial tasks. Tasks that were of medium level of difficulty had the best results in the monoscopic condition. There were also some insights in terms of age, and gender. Females performed better in the stereoscopic condition than the monoscopic one. Also, participants over 25 years also had a statistically significance higher performance with the stereoscopic condition.

The performance of the stereoscopic condition was not highest under every condition, which may be due to the animation of the visual representation. By adding animations, participants had different points of view from which to base their relative object positions.

8.7 Future Work

The results of this study gave initial insights into cases where stereoscopic applications would be of benefit to a user. Followup studies focusing on stereopsis related to demographics such as gender, age. Additionally future studies need to confirm the distances where stereoscopic applications and animation offer the best benefits.

CHAPTER 9. CONCLUSION

9.1 Summary of Hypothesis

Most of hypothesis have been addressed, supported or rejected through the three studies conducted. They are summarized below.

9.1.1 H1: Adding stereopsis as a depth cue will have a positive effect on performance of a visual-spatial task.

This hypothesis was confirmed during the three studies. Overall the stereoscopic condition yielded a higher performance than the monoscopic condition across different tasks, cases, volume orientations, and animation. However, the overall benefit of the stereoscopic condition depended on the difficulty of the task and the orientation of the volume. Future work needs to be conducted to understand the relationships between these variables. In general stereoscopic conditions offer a benefit for visual spatial tasks over monoscopic condition.

9.1.2 H2: Performance will decrease as the distance between objects being judged decreases. The stereoscopic condition will yield higher performances in the most "difficult" cases (i.e. smaller distances).

This hypothesis was partially confirmed. Performance did decrease when the difficulty of the task increased. For distances beyond 15 mm the effect of the stereoscopic condition over the monoscopic condition was not very pronounced. For distances between 5 mm and 15 mm the stereoscopic condition had the largest benefit. For distances below 5mm there were minor benefits found for the stereoscopic condition.

This was consistent for the three studies. This can be an indication of what distances offer the most benefit for visual spatial tasks. This can have an implication for designing stereoscopic applications: the stereoscopic condition can be turn on automatically at distances between 5 mm and 15 mm, where it offers the maximum benefit. It can be turn off at the other distances. Turning on the stereoscopic conditions only when necessary could minimize other human factors effects such as fatigue and headaches.

9.1.3 H3: Different object orientations will have an effect on performance. Orientations at 45^o degree angles will result in decreased performance. The stereoscopic condition will have higher performance than the monoscopic condition at these angles.

This hypothesis was also confirmed. This was only tested in the first and the second studies. The down view was the easiest because participants had a top down orientation to assess the relative position of the objects, making it almost a 2D problem. The 45^o angle views were the hardest for the participants. These produced artifacts that can create inaccurate impressions of the positions of objects in a truly 3D setting. The stereoscopic condition had higher performance for most of the object orientations, especially the most difficult views. The stereoscopic condition may offer a benefit over the monoscopic condition over angled object views.

9.1.4 H4: Animation will have a positive effect on performance for the monoscopic conditions as it allows different views of the representation before the task is performed.

Overall the stereoscopic condition yielded better results than the monoscopic condition for the animated tasks. However because the animation provides participants with multiple points of view it gives them additional cues to assess the relative position of the objects. The overall benefits of the stereoscopic condition were decreased because of the animation.

9.1.5 H5: Demographics such as age and gender will have an effect on performance of visual-spatial tasks. The stereoscopic condition will have an overall positive effect across different demographic groups.

This hypothesis was partially confirmed. There was an interaction between the performance, display condition, and the age group. Different demographics such as age had different performances. The stereoscopic condition had an overall positive effect for both age groups in the second study, and a positive effect for participants of 25 and over for the third study. These interaction effects need to be studied with large sample sizes in future studies. In order to understand how individual differences affect performance in stereoscopic applications.

9.1.6 H6: Higher performance in paper visual spatial tasks will correlate with higher performance on the computer visual spatial task. The stereoscopic condition will have an overall greater effect on lower performing individuals than the monoscopic condition.

This hypothesis was rejected. In all three studies there was no relationship between the mental rotation tasks, and the paper rotation tests to the visual spatial task performance. This is contrary to what literature has found between visual spatial tasks and surgical skills [55, 60]. An explanation of this difference is that previous studies have studied the correlation of performance between surgical tasks and the paper tasks, and surgical tasks include navigation, grasping, and relative position judgment tasks, while the studies on this work only studied relative position judgment tasks.

9.2 Summary

In general stereopsis offers a benefit in performance for visual spatial tasks in a medical context. In previous studies stereoscopic applications were found to yield mixed results, depending on the type of task, and domain. While stereoscopic conditions of some tasks such as the manipulation of objects have shown positive results [29], there were some basic research questions on the benefits of stereoscopic applications for visual spatial tasks. Visual spatial tasks are needed in medicine to understand the relationships between shapes and organs, for navigation, and surgery. The visual spatial task performed in the three studies confirmed that the stereoscopic condition provided an overall benefit in performance.

9.3 Future Work

Even though the patterns show that stereoscopic applications offer a benefit over monoscopic applications, there are many questions that need to be researched. One question that remains is the relationship between demographics and the benefits of the stereoscopic applications in visual spatial tasks.

Other areas that need to be studied are the individual human factors and performance. Stereoscopic applications can create visual fatigue, nausea, and eye strain. How would these human factors affect task performance? Future work needs to integrate stereoscopic applications and human factors.

Additionally other tasks beyond visual spatial tasks such as navigation, and object recognition in medical contexts need to be studied.

9.4 Acknowledgements

There were many involved that supported through my graduate education and this dissertation. First I would like to thank my partner, Riley Hostetter, for their continuous support, all the late nights and cancelled plans. I would also like to thank my parents, Manuel Martinez, and Blanca Escobar who kept motivating me about grad work.

I would like to thank my advisor Dr. Eliot Winer, Eliot took a chance on me and helped me believe in myself and take chances that I wouldn’t have. I wouldn’t be where I am today without his guidance.

Finally I would like to thank Dr. Eric Foo and Dr. Ken Kopecky, my first colleagues during the grad school adventure. We spent many late nights in the lab, and I will cherish those times.

APPENDIX. PRE-SURVEY QUESTIONS

Age: ______

Sex: Male / Female

Job title (Years of experience if applicable) ______________________ (______)

Which anatomy classes have you completed?

Which anatomy classes are you currently taking?

Do you have any experience with 3D visualization? (3D movies/games) Yes / No

(Please elaborate)

Do you have any experience with any medical visualization software? Yes / No

(Please elaborate)

Do you have any experience with Medical Reconstruction/Visualization? Yes / No

(Please elaborate)

Do you have any experience with Mental Rotation Testing? Yes / No

(Please elaborate)

Do you have any experience with Paper Folding Testing? Yes / No

(Please elaborate)

REFERENCES

1. O'Connor, A.R., et al., *The functional significance of stereopsis.* Investigative ophthalmology & visual science, 2009. **51**(4): p. 2019-2023.

2. Melmoth, D.R. and S. Grant, *Advantages of binocular vision for the control of reaching and grasping.* Experimental Brain Research, 2006. **171**(3): p. 371-388

3. Melmoth, D.R., et al., *Grasping deficits and adaptations in adults with stereo vision losses.* Investigative ophthalmology & visual science, 2009. **50**(8): p. 3711-3720.

4. Held, R.T. and T.T. Hui, *A guide to stereoscopic 3D displays in medicine.* Academic radiology, 2011. **18**(8): p. 1035-1048.

5. de Vries, S.C. and P. Padmos. *Steering a simulated Unmanned Aerial Vehicle using a head-slaved camera and HMD*. 1997. International Society for Optics and Photonics.

6. Jayaram, S., et al., *Assessment of VR technology and its applications to engineering problems.* Journal of Computing and Information Science in Engineering, 2001. **1**(1): p. 72-83.

7. Van Beurden, M., W.A. Ijsselsteijn, and J.F. Juola, *Effectiveness of stereoscopic displays in medicine: a review.* 3D Research, 2012. **3**(1): p. 1-13

8. Andel, C., et al., *The economics of health care quality and medical errors.* Journal of health care finance, 2012. **39**(1): p. 39.

9. Galloway, R., and Terry Peters, *Overview and History of Image-Guided Interventions*, in *Image-Guided Intervention*, T.M. Peters, Editor. 2008, Springer.

10. Lipinski, J.K., *Some Observations on Early Diagnostic Radiology in Canada.* Canadian Medical Association, 1983. **129**(7): p. 766-768.

11. *Imaginis: Brief History of CT*. 2007 [cited 2008 October 22]; Available from: http://www.imaginis.com/ct-scan/history.asp.

12. Hornak, J. *The Basics of MRI*. [cited 2008 22 october 2008]; Available from: http://www.cis.rit.edu/htbooks/mri/inside.htm.

13. Elvins, T.T., *A survey of algorithms for volume visualization.* SIGGRAPH Comput. Graph., 1992. **26**(3): p. 194-201.

14. Parker, S., et al., *Interactive ray tracing*, in *Proceedings of the 1999 symposium on Interactive 3D graphics*. 1999, ACM: Atlanta, Georgia, United States. p. 119-126.

15. Huang, J., et al. *A Practical Evaluation of Popular Volume Rendering Algorithms*. in *Proceedings of the 2000 IEEE symposium on Volume visualization*. 2000.

16. Bankman, I.N. and S. Morcovescu, *Handbook of Medical Imaging. Processing and Analysis.* Medical Physics, 2002. **29**(1): p. 107-107.

17. Rosset, A., L. Spadola, and O. Ratib, *OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images.* Journal of Digital Imaging, 2004. **17**(3): p. 205-216.

18. Luursema, J.-M., et al., *Optimizing conditions for computer-assisted anatomical learning.* Interacting with Computers, 2006. **18**(5): p. 1123-1138.

19. Ahmed, J., *The sensitivity and specificity of nonmydriatic digital stereoscopic retinal imaging in detecting diabetic retinopathy.* Diabetes care, 2006. **29**(10): p. 2205.

20. Association, A.O. *New Hand-Held 3D Gaming Devices May Help Uncover Undiagnosed Vision Problems*. 2011 [cited 2012; Available from: http://www.aoa.org/x17309.xml.

21. Lambooij, et al., *Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review.* Journal of Imaging Science and Technology, 2009. **53**(3): p. 30201.

22. Patterson, R., *Review paper: human factors of stereo displays: an update.* Journal of the Society for Information Display, 2009. **17**(12): p. 987-996.

23. Lambooij, M.T.M., W.A. Ijsselsteijn, and I. Heynderickx. *Visual discomfort in stereoscopic displays: a review*. in *Stereoscopic Displays and Virtual Reality Systems XIV*. 2007. San Jose, CA, United States: International Society for Optics and Photonics.

24. Patey, N. *3D is dangerous/not dangerous: Nintendo 3DS warning label edition*. 2010 [cited 2012; Available from: http://www.engadget.com/2010/12/29/3d-is-dangerous-not-dangerous-nintendo-3ds-warning-label-edit/.

25. Patterson, R., L. Moe, and T. Hewitt, *Factors that affect depth perception in stereoscopic displays.* Human factors, 1992. **34**(6): p. 655-667.

26. Patterson, R., et al., *Binocular rivalry and head-worn displays.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2007. **49**(6): p. 1083-1096.

27. Meesters, L.M.J., W.A. Ijsselsteijn, and P.J.H. Seuntiens, *A survey of perceptual evaluations and requirements of three-dimensional TV*. Circuits and Systems for Video Technology, IEEE Transactions on, 2004. **14**(3): p. 381-391.

28. Mayer, U., et al. *Is eye damage caused by stereoscopic displays?* 2000. International Society for Optics and Photonics.

29. McIntire, J.P., P.R. Havig, and E.E. Geiselman, *Stereoscopic 3D displays and human performance: a comprehensive review*. Displays, 2014. **35**(1): p. 18-26.

30. Risucci, D.A., *Visual spatial perception and surgical competence.* The American Journal of Surgery, 2002. **184**(3): p. 291-295.

31. Keehner, M., et al., *Learning a spatial skill for surgery: how the contributions of abilities change with practice.* Applied Cognitive Psychology, 2006. **20**(4): p. 487-503.

32. Epstein, W. and S. Rogers, *Perception of space and motion*. 1995: Academic Press.

33. Cutting, J.E., *Potency, and Contextual Use of Different Information about Depth.* Perception of space and motion, 1995: p. 69.

34. Chen, J., W. Cranton, and M. Fihn, *Handbook of visual display technology*. 2012: Springer.

35. Holliman, N.S., et al., *Three-dimensional displays: a review and applications analysis.* Broadcasting, IEEE Transactions on, 2011. **57**(2): p. 362-371.

36. *3D Display Technology and Market Forecast Report*. 2010, DisplaySearch: Austin TX.

37. Beurden, v., et al., *Effectiveness of stereoscopic displays in medicine: A review.* 2012. **3**(1): p. 13.

38. Getty, D.J., R.M. Pickett, and C.J. D'Orsi, *Stereoscopic digital mammography: improving detection and diagnosis of breast cancer.* International Congress Series, 2001. **1230**(0): p. 538-544.

39. Kickuth, R., et al., *Stereoscopic 3D CT vs standard 3D CT in the classification of acetabular fractures: an experimental study.* 2002. **75**(893): p. 422-427.

40. Yaniv, Z. and K. Cleary, *Image-guided procedures: A review.* Computer Aided Interventions and Medical Robotics, 2006. **3**.

41. Reitinger, B., *Liver surgery planning using virtual reality.* IEEE Computer Graphics and Applications, 2006. **26**(6): p. 36.

42. Aziz, O., et al., *Laparoscopic Versus Open Surgery for Rectal Cancer: A Meta-Analysis.* Annals of Surgical Oncology, 2006. **13**(3): p. 413-424.

43. Votanopoulos, K., *Impact of three-dimensional vision in laparoscopic training.* World journal of surgery, 2008. **32**(1): p. 110.

44. Perkins, N., *Learning to use minimal access surgical instruments and 2-dimensional remote visual feedback: How difficult is the task for novices?* Advances in health sciences education, 2002. **7**(2): p. 117.

45. Crosthwaite, G., et al., *Comparison of direct vision and electronic two- and three-dimensional display systems on surgical task efficiency in endoscopic surgery.* British Journal of Surgery, 1995. **82**(6): p. 849-851.

46. Hanna, G.B., S.M. Shimi, and A. Cuschieri, *Randomised study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy.* The Lancet, 1998. **351**(9098): p. 248-251.

47. John, M., et al., *Tactical Routing Using Two-Dimensional and Three-Dimensional Views of Terrain.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2001. **45**(1): p. 1409-1413.

48. Tory, M., et al. *Eyegaze analysis of displays with combined 2D and 3D views*. in *Visualization, 2005. VIS 05. IEEE*. 2005.

49. Tory, M., et al. *Combining 2D and 3D views for orientation and relative position tasks*. in *Proceedings of the SIGCHI Conference on Human Factors in Computing System*. 2004. ACM.

50. Smallman, H.S., *Information availability in 2D and 3D displays.* IEEE Computer Graphics and Applications, 2001. **21**(5): p. 51.

51. John, M.S., et al., *The use of 2D and 3D displays for shape-understanding versus relative-position tasks.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2001. **43**(1): p. 79-98.

52. Sando, T., M. Tory, and P. Irani, *Effects of animation, user-controlled interactions, and multiple static views in understanding 3D structures*, in *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*. 2009, ACM: Chania, Crete, Greece. p. 69-76.

53. Peters, M., et al., *A Redrawn Vandenberg and Kuse Mental Rotations Test - Different Versions and Factors That Affect Performance.* Brain and Cognition, 1995. **28**(1): p. 39-58.

54. Ekstrom, R.B. and J.W. French, *Manual for kit of Factor-Referenced Cognitive Tests*, ed. H.H. Harman. 1976: Educational Testing Service.

55. Wanzel, K.R., et al., *Effect of visual-spatial ability on learning of spatially-complex surgical skills.* The Lancet, 2002. **359**(9302): p. 230-231.

56. Foo, J.L., T. Lobe, and E. Winer. *Isis: Patient Data Visualization and Surgical Planning in an Interactive Virtual Environment*. in *ASME-AFM 2009 World Conference on Innovative Virtual Reality*. 2009. American Society of Mechanical Engineers.

57. Hill, F.S., *Computer Graphics Using Open GL*. 2000: Prentice Hall.

58. Hegarty, M., et al., *The role of spatial cognition in medicine: Applications for selecting and training professionals.* Applied spatial cognition, 2007: p. 285-315.

59. Getty, D.J. and P.J. Green, *Clinical applications for stereoscopic 3‐ D displays.* Journal of the Society for Information Display, 2007. **15**(6): p. 377-384.

60. Wanzel, K.R., et al., *Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance.* Surgery, 2003. **134**(5): p. 750-757.

61. *Diagnostic Errors More Common, Costly And Harmful Than Treatment Mistakes*. 2013 [cited 2014 December 11].

62. Wachter, R.M., *Why diagnostic errors don't get any respect—and what can be done about them.* Health Affairs, 2010. **29**(9): p. 1605-1610.

63. Ahmed, J., et al., *The sensitivity and specificity of nonmydriatic digital stereoscopic retinal imaging in detecting diabetic retinopathy.* Diabetes care, 2006. **29**(10): p. 2205-2209.

64. Stewart, N., et al., *Stereoscopy in diagnostic radiology and procedure planning: Does stereoscopic assessment of volume‐ rendered CT angiograms lead to more accurate characterisation of cerebral aneurysms compared with traditional monoscopic viewing?* Journal of medical imaging and radiation oncology, 2014. **58**(2): p. 172-182.

65. D'Orsi, C.J., et al., *Stereoscopic digital mammography: Improved specificity and reduced rate of recall in a prospective clinical trial.* Radiology, 2013. **266**(1): p. 81-88.

66. Qiu, T.-m., et al., *Virtual reality presurgical planning for cerebral gliomas adjacent to motor pathways in an integrated 3-D stereoscopic visualization of structural MRI and DTI tractography.* Acta neurochirurgica, 2010. **152**(11): p. 1847-1857.

67. Mistry, M., V.A. Roach, and T.D. Wilson, *Application of stereoscopic visualization on surgical skill acquisition in novices.* Journal of surgical education, 2013. **70**(5): p. 563-570.

68. Martinez-Escobar, M., et al. *Assessment of visual-spatial skills in medical context tasks when using monoscopic and stereoscopic visualization*. in *SPIE Medical Imaging*. 2013. International Society for Optics and Photonics.

69. Taffinder, N., et al., *The effect of a second-generation 3D endoscope on the laparoscopic precision of novices and experienced surgeons.* Surgical endoscopy, 1999. **13**(11): p. 1087-1092.