12-2007

RADIAL OUTFLOW IN TELEOPERATION: A POSSIBLE SOLUTION FOR IMPROVING DEPTH PERCEPTION

Joshua Gomer
Clemson University, jagomer@gmail.com

RADIAL OUTFLOW IN TELEOPERATION:
A POSSIBLE SOLUTION FOR IMPROVING DEPTH PERCEPTION

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Applied Psychology

by
Joshua A. Gomer
December 2007

Accepted by:
Dr. Christopher Pagano, Committee Chair
Dr. Rick Tyrrell
Dr. Ian Walker

ABSTRACT

Practical experience has shown that operators of remote robotic systems have difficulty perceiving aspects of remotely operated robots and their environments (e.g. Casper & Murphy, 2003). Operators often find it difficult, for example, to perceive accurately the distances and sizes of remote objects. Past research has demonstrated that employing a moveable camera that provides the operator optical motion allows for the perception of distance in the absence of other information about depth (Dash, 2004). In this experiment a camera was constrained to move only forward and backward, thus adding monocular radial outflow to the video stream. The ability of remote operators to perceive the sizes of remote objects and to position a mobile robot at specific distances relative to the object was tested. Two different conditions were investigated. In one condition a dynamic camera provided radial outflow by moving forward and backward while atop a mobile robot. In the second condition the camera remained stationary atop the mobile robot. Results indicated no differences between camera conditions, but superior performance for distance perception was observed when compared to previous research (Dash, 2004). This thesis provides evidence that teleoperators of a terrestrial robot are able to determine egocentric depth in a remote environment when sufficient movement of the robot is involved.

DEDICATION

This thesis is dedicated to my lovely wife Kerry. It was a difficult decision for us to leave our families behind and begin the graduate school adventure in such an unfamiliar location. And the last two years have definitely not been without adversity. However, considering our many new friends and accomplishments, as well as a deeper appreciation for each other, I am very thankful you supported our decision. Your love and support motivates me to continue to better myself, and for all that you are I love you dearly.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Chris Pagano for advising this work, along with Dr. Rick Tyrrell and Dr. Ian Walker for their valuable contributions as committee members. I consider myself fortunate to have worked with each of you. I also extend my gratitude to Lamar Durham and the Clemson Physics Machine Shop for providing our Lab with the necessary modifications to our robot. In addition, I will always greatly appreciate my colleagues at Clemson for their willing, continuous, and valuable feedback. Finally, I thank Barbara Ramirez, for without her expertise, much of this would have just sounded silly.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

A Brief History of Applied Teleoperation

At the beginning of World War II, engineers started to develop robotic systems to replace humans for increased efficiency and faster completion of repetitive industrial production tasks (Stassen & Smets, 1997). These primitive robots were successfully fabricated and programmed to accomplish rote tasks; however, they "failed immediately if the task deviated from the specific constraints for which the robot was built" (Smets, 1995, pg. 182). Over the next three decades, advances in industrial robotics continued to show potential for an increased variety of applications. While valuable for industrial increases in efficiency, precision, and speed, these robots also began to demonstrate potential in dangerous work environments which were unsafe for humans.

In the 1970's, the real-time control of robots through the use of remote control devices began with undersea applications (Smets, 1995). This newer style of machine was used for approximately 20 years before this remote, or teleoperator control, was operationally defined by Sheridan (1992) as "the extension of human sensing and manipulation capabilities by coupling it to remote artificial sensors and actuators" (pg. 393). In this context, the human operation of robots involves various types of sensors and communication devices (Sheridan, 1989). Teleoperated robots encompass a diversity of forms, including terrestrial wheeled and tracked vehicles, quadrupeds, snakes and other climbers, unmanned aerial vehicles (UAVs), and aquatic submersibles. Many times the primary difference between these more modern hybrid systems and industrial robots is the introduction of a novel action or movement (Smets, 1995).

The primary use of these remote controlled devices involves situations inadvisable for humans. To minimize human casualties and maximize safety, robots are sent into hazardous environments to assess dangerous situations in both military and civilian applications. Specific examples of dangerous teleoperation tasks include UAV aerial military intelligence collection, weather surveillance, and zero gravity space repair. Other hazardous applications include inspection and maintenance of offshore oil platforms (Lin & Kuo, 1999), as well as nautical archaeology and marine geology data collection (Negahdaripour & Madjidi, 2003) and undersea wreckage investigation. Unsafe terrestrial activities include toxic waste detection and cleanup, sewer inspection, mining, disaster recovery, and handling of radioactive materials during surgery (see Mailhot, 1996).

One of the most important uses of teleoperated robots is in urban search and rescue (USAR), first seen on September 11, 2001 (Casper & Murphy, 2003; Murphy, 2004). In potentially volatile remote locations, such as the hot zones encountered at Ground Zero, robots provide a viable means of exploration. These robots, however, are only useful when capable of providing an operator with accurate information about the identity and condition of objects, structures, and human casualties (Woods et al., 2004). Arriving within 6 hours of the attack of the World Trade Center onset, the Center for Robot-Assisted Search and Rescue (CRASAR) joined the effort to search through the rubble, with eight deployments being undertaken in the days after 9/11. During these deployments robots were occluded from direct line of sight an average of 18% of the time, the degraded conditions adding to these already compromised perceptions of the remote environment.

Experience at the World Trade Center and other disaster sites has revealed that for teleoperated robots to be effective, many challenges must be met, including unnatural perspectives, variable lighting, degraded optical information due to smoke and dust, lack of information about the state of the robot (Burke et al., 2004), and the loss of depth information with a flat screen. According to Murphy (2004), one of the most significant problems at Ground Zero "was the lack of depth perception" (pg. 57). These deployments also revealed poor user interfaces, infrared devices rendered useless by extreme heat, the superiority of color cameras in survivor searches, and the superiority of black and white, high resolution cameras in structural assessment. These difficult conditions resulted in the video footage recorded at the World Trade Center revealing previously unseen objects even five months after the disaster (Murphy, 2004). Since September 11, 2001, robots have been more widely accepted by the rescue community for victim searches, determining safe and efficient paths through rubble, structural inspection, and hazardous material detection (Murphy, 2004).

Since teleoperation involves an operator, mental and physical fatigue also limit the usefulness of these robots. For example, mental fatigue due to a lack of sleep is common during disasters. Burke (2004) stated that "it is conventional wisdom that a responder will get less than 3 hours of continuous sleep during the first 48 hours of an incident" (pg. 89). Operators must be able to overcome their own mental fatigue during a disaster to be able to effectively operate a robot in degraded conditions. Physical fatigue due to lack of sleep and a heightened level of activity must also be overcome. Many times robots must be carried or man-packed until the actual disaster site is reached, due to a lack of traversable roads.

Theoretical Considerations of Teleoperation

In addition to the applied problems exemplified at Ground Zero, there are also many theoretical issues with teleoperation. Robin Murphy, one of the world's leading roboticists, stated that "progress in mobile robots relies on progress in perception" (1999, pg. 105). These perceptual issues include the definition of telepresence along with the contrasting concepts of its maximization and functional presence, destructive mapping, the decoupling of perception and action, binocular camera systems, monocular visual cues, and optical motion.

Telepresence refers to achieving a state of presence without actually being present. For example, if a robot and the teleoperator are in different rooms, the controller must rely on the robotic sensors experiencing a mediated level of presence. According to Sheridan (1989), this telepresence is achieved only if information about the robot and the remote environment is communicated as realistically as possible to the teleoperator. The figure based on the Barometer of Presence Fidelity created in the late 1990s by Agah and Tanie (1999) seen in Figure 1 illustrates the relationship between 4 levels of mediated presence and reality:

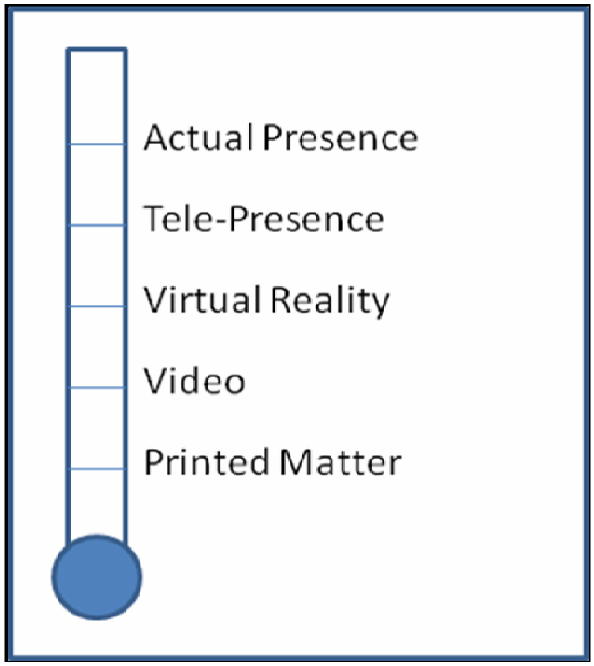

Figure 1: A Barometer for Presence Fidelity (after Agah & Tanie, 1999, pg. 106).

As this continuum suggests, telepresence brings "the remote user an experience as close as possible to actual presence" (Agah & Tanie, 1999, pg. 106). This degree of realism is achieved due to the live presentation of a real environment instead of the simulated worlds found in the remaining three levels of perception. Compared to the limitations in software or gaming programs, for example, telepresence lends itself to increased "dynamic possibilities" and the observation of "actual happenings" (Agah & Tanie, 1999, pg. 105).

The higher the level of telepresence the more realistic is the experience of the teleoperator. As a result, the amount of technology necessary is increased, including such equipment as infrared sensors, multiple cameras, machine vision and autonomous obstacle detection. Although these sensors can fail as evidenced by Ground Zero, they demonstrate potential in teleoperation.

An alternative to maximizing telepresence is to emphasize functional presence. In this situation, according to Tittle et al., a teleoperator has enough information available to their senses to be able to function effectively (2002). This level does not attempt to duplicate reality; rather it focuses on the ability to perceive only the relevant information in a remote environment. As current technology does not allow for the duplication of a remote environment through the sensors of a robot, most research in this area focuses on increasing functional presence.

Destructive mapping refers to the loss of information that occurs when the 3-D world is mapped onto a 2-D image captured on the retina of the eye. Even though the retina is not capable of recording the world in three dimensions, the brain, for the most part, replaces the dimension that is lost. Theories of perception influenced by cognition, also referred to as indirect perception, argue that this destructive mapping, or loss of information, naturally occurs between the proximal stimulus, the eye, and the distal stimulus, the object, being perceived. Other theories of perception, such as ecological, or direct perception, suggest that this loss is less an issue of cognition and more one of the amount of information presented to a perceiver. For example, as perceivers move about in an environment, they are able to gather more information from it. Replacing a perceiver with a camera, as in teleoperation, increases the level of destructive mapping since it further limits the potential for information gathering. Increasing the number of cameras and their capabilities does not necessarily eliminate this remote perception problem (Tittle et al., 2002).

In teleoperation, visual information is mediated by a camera and a monitor, removing an operator physically from an environment. Decoupling perception and action occurs and this breaks the natural "perception-action link" (Tittle et al., 2002, pg. 260), since perception is normally tightly coupled to action (see Bingham & Pagano, 1998; Pagano & Bingham, 1998). The resulting mediated perception requires operators to act on interpretations instead of first-hand information, diminishing the amount of information available to the perceiver, in particular, "reliable estimates of depth can only be made if the 'poor' stimulus is 'enriched' for unambiguous interpretation by some cognitive inference" (Smets et al., 1987, pg. 1032). This decoupling also leads to the loss of vestibular feedback and proprioceptive cues, creating sensory conflicts for teleoperators. Physiologically, these systems normally provide information regarding the rate of change in velocity and in the orientation of the body with respect to gravity. Since teleoperators are in a secondary location, they will receive no accurate vestibular or proprioceptive feedback about the location of the robot. For some applications teleoperation may be performed in uncoupled motion, meaning the operator is moving while operating a mobile robot. The complexity of this situation further exacerbates this decoupling, as it generates two conflicting patterns of motion in different directions and at different velocities, many times resulting in sickness.

Stereoscopic cameras are frequently integrated to mimic the human visual system in teleoperation, but this usually increases the complexity of a difficult situation for the designer and the operator. Stereoscopic systems require higher resolution cameras, necessitating higher-bandwidth and power, fed to a head-mounted display, in turn limiting the operator's field of view. As a result, the "variable relative position, tilt, and rotation" of cameras must translate in synch with the angle of rotation of the head and eyes of the teleoperator (Agah & Tanie, 1999, pg. 110), increasing the complexity of software and hardware. Stereoscopic displays also result in user fatigue, discomfort, and acclimatization after long periods (see Agah & Tanie, 1999). For example, it takes longer to achieve head-mounted stereoscopic depth when compared to viewing a monitor, "important if critical stages of a remote operation require switching between normal and stereoscopic viewing" (Reinhardt-Rutland, 1996, pg. 244).

One method of addressing the disadvantages of stereoscopic camera systems is to take advantage of monocular vision. Binocular disparity is lost without differences provided by two laterally spaced eyes and depth perception suffers as a consequence (e.g. Woods, et al., 2004; Reinhardt-Rutland, 1996). However, this situation can be mitigated by monocular vision utilizing linear perspective, interposition, light, shadow, and static pictorial cues (e.g. Goldstein, 2007). This approach eliminates the differences resulting from two cameras, or eyes, in the case of humans, referred to as binocular disparity. According to Reinhardt-Rutland (1996, pg. 242), these perceptual cues will more often than not "remain broadly intact in remote operation"; however, pictorial information should be enhanced so that the remote perceiver is not misled. Linear perspective cues, for example the converging of parallel lines (e.g. train tracks) against the horizon, in conjunction with interposition cues, can provide size and distance information, by the blocking of objects more distant from an observer by those closer. Light and shadow primarily provide shape cues in monocular vision. Static pictorial cues include edges, relative visual size, and orientation-in-depth. Indicating a discontinuity in depth, an edge is an important feature of objects in the environment in navigating or perceiving our surroundings. Visual size helps a perceiver relate the sizes and distances of various objects both to one another and the operator. In addition, understanding the orientation in depth of an object can assist an observer in determining the true shape of an object. For example, a rectangle may be perceived incorrectly as a trapezoid depending on its orientation in depth. Improving the operator's ability to utilize these classes of perceptual information can be an alternative to the complexities and disadvantages of a binocular system.

One of the most important factors in teleoperation is optical motion, specifically motion information, which is generated by motion parallax or the systematic visual motions on the retina by objects in the visual field of a moving observer. This parallax provides rates of visual motion greater for near objects than for more distant ones. In other words, for a given rate of motion, optic flow reveals faster expansion and contraction of objects closer to the observer than objects far away. This optical motion, or optic flow, provides information to a perceiver about the environment, as well as the objects, action possibilities, and rate of movement in it. For example, the change in an object's size or shape based on the movement of a perceiver provides additional cues due to the resulting change in perceptual point of reference. Gibson (1979) argued that when experienced directly, the best way for a perceiver to gain information regarding an environment would be to increase locomotion in it, thus increasing optic flow. This philosophy has influenced several recent research studies.

Recent Teleoperation Research

A recent study focusing on improving functional presence in teleoperation involved a design by Agah and Tanie (1999). This study investigated the use of sonar feedback combined with a decoupled camera, or a camera capable of movement independent of the movement of the robot. In their design, teleoperators navigated a mobile robot 5.5 cm in diameter and 3 cm tall, outfitted with a pan and tilt camera, through two scaled models of a museum. Exhibits in 3-D were provided as viewing targets and movements of the robot were used to achieve camera rotation, as well as forward and back camera motion. Automatic focus was disabled in this study, controlling for target distance information. In demonstrating that their multi-media interface and control system can be used for operator exploration of a remote site, Agah and Tanie (1999) explored stereoscopic camera system design problems and the benefit of having a camera capable of independent movement of the robot. In particular, they focused on the augmentation of visual information with sonar information as a fusion between the operator and the robot, defining such "intelligence fusion" as the integration of the human user's intelligence with the robot's (Agah & Tanie, 1999, pg. 105). While their results support semi-autonomous robotics for increased teleoperation performance, their findings concerning the limiting factors of stereovision and the benefit of using a decoupled camera system are most important to the current study.

More recently, Sekmen at al. (2003) further examined sonar obstacle detection with semi-autonomous robots, investigating four conditions of teleoperator location and prior knowledge of navigational course. Participants were given a short familiarization session and a decoupled robot equipped with a pan and tilt camera. They then took part in four trials in which they were instructed to move the robot from the same starting location to differing ending target locations. However, sonar readings sent from the robot were needed to assist 84% of the operators with collision avoidance, a situation attributed "to the fact that the camera images do not provide enough depth information about the objects near the robot" (Sekmen et al., 2003, pg. 15). These results illustrate the difficulties with depth perception in teleoperation, providing additional support for decoupled camera systems, as well as increasing system complexity to improve depth perception.

At approximately the same time, a third study conducted by Hughes et al. (2003), also investigated the utility of providing decoupled cameras to the teleoperator, using them "to mitigate some of the problems with situational awareness, and increase the effectiveness of search tasks" (pg. 1). In addition, a decoupled camera, operating independent of robotic movement, has the advantage of providing a natural mapping (walking forward and peering right). However, these researchers hypothesized that switching between teleoperation subtasks of navigation and inspection with a decoupled camera could create difficulties in situation awareness for teleoperators. The results of their study indicate that a camera mounted on a robot, dependent on its movements increased the cognitive burden of mediating perception in relation to the robot, a hypothesis supported by their study. This fixed camera required repeated physical adjustment, thereby increasing power usage and the instances that the robot was stuck or obstructed. A camera independent of robotic movement has the disadvantage of increasing the degrees of freedom controlled by the operator. Even with these disadvantages, their results indicated that providing a camera controlled independently of mobile robot orientation provides significant benefits to functional presence in search activities.

Radial Outflow from Optical Motion

As these remote perception studies indicate, the use of a decoupled camera in teleoperation merits further study. The theoretical basis for such an investigation involves understanding the importance of optical motion. One of the first studies in this area was conducted by Rogers and Graham (1979), who investigated the role of motion parallax in human depth perception. Subjects in their study viewed monocular random dot stereograms having a simulated 3-D structure revealed only through optical motions. Under one condition, the optical motions were coupled to the side-to-side movements of the subject's head, and under the second the optical motions were passive. Their results suggest that whether self-produced or passively produced, "motion parallax can be a sufficient cue to the shape and depth of three-dimensional surfaces, in the absence of all other depth cues" (Rogers & Graham, 1979, pg. 132). This finding underscores the importance of increasing the amount of camera motion available to an operator, especially in search and rescue operations where other sources of information may not be reliable.

Bingham and colleagues have generalized these findings to the use of forward and back head motions (see Bingham & Stassen, 1994; Bingham & Pagano, 1998; Pagano & Bingham, 1998). Such motions create radial outflow as the participant moves toward a target. Previous research had shown that this radial outflow contains information about the observer's time-to-contact if the motion continues at a constant velocity. This optical time-to-contact, specified by the relative rate of dilation of the target's image is referred to as τ, or *tau* (Lee, 1974). Bingham and Stassen (1994) demonstrated that tau can provide information about egocentric depth if it is assumed that the head moves toward and away from the target with a consistent sinusoidal velocity profile:

$$\mathbf{D = \tau\ 2\pi A/P} \qquad \textbf{(1)}$$

where D equals the distance to the target, τ the optically specified time-to-contact at peak velocity, A the amplitude of head motion, and P the period of the head motion, thus demonstrating mathematically that depth information is conveyed by radial outflow (for other mathematical derivations see Bruckstein et al., 2005; Bingham & Pagano, 1998). Bingham and Pagano (1998) later demonstrated that this information can be perceived by human observers through their study using a helmet-mounted camera and display to isolate the optic flow created by the participants' voluntary head movements toward a target.

Investigating Radial Outflow in Teleoperation

Dash (2004) extended the work of Bingham by examining self-produced radial outflow as providing effective information about egocentric depth in teleoperation. As seen in Figure 2, participants in this study viewed white squares against black space under three different conditions. Targets in this experiment consisted of white foam board squares at 5, 11, and 14 degrees of visual angle from the camera lens, producing three sizes on the monitor of 7.6, 16.5, and 21.3 cm respectively, thus controlling for visual angle as seen in Figure 3. The five target distances from the initial position of the camera were 75, 100, 125, 150, and 175 cm.

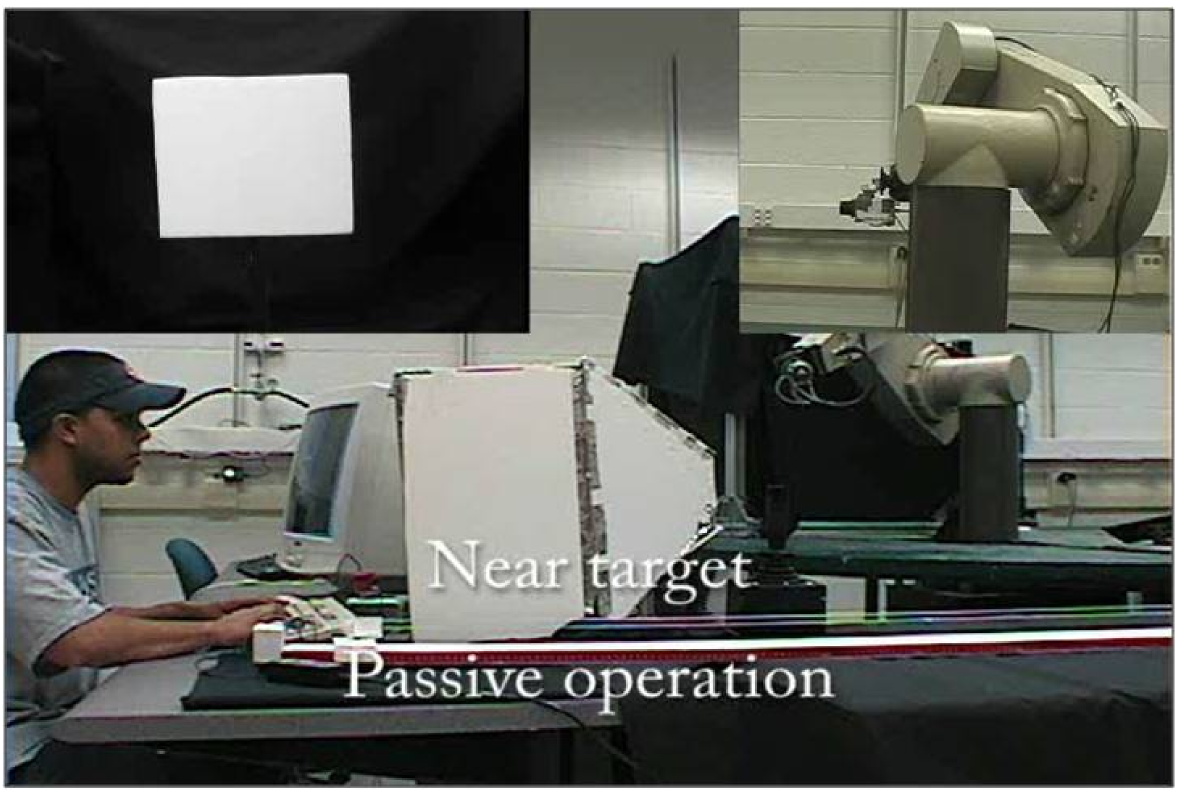

Figure 2: Three images from Dash (2004) of a white square against black space, the robotic Puma arm, and the participant setup.

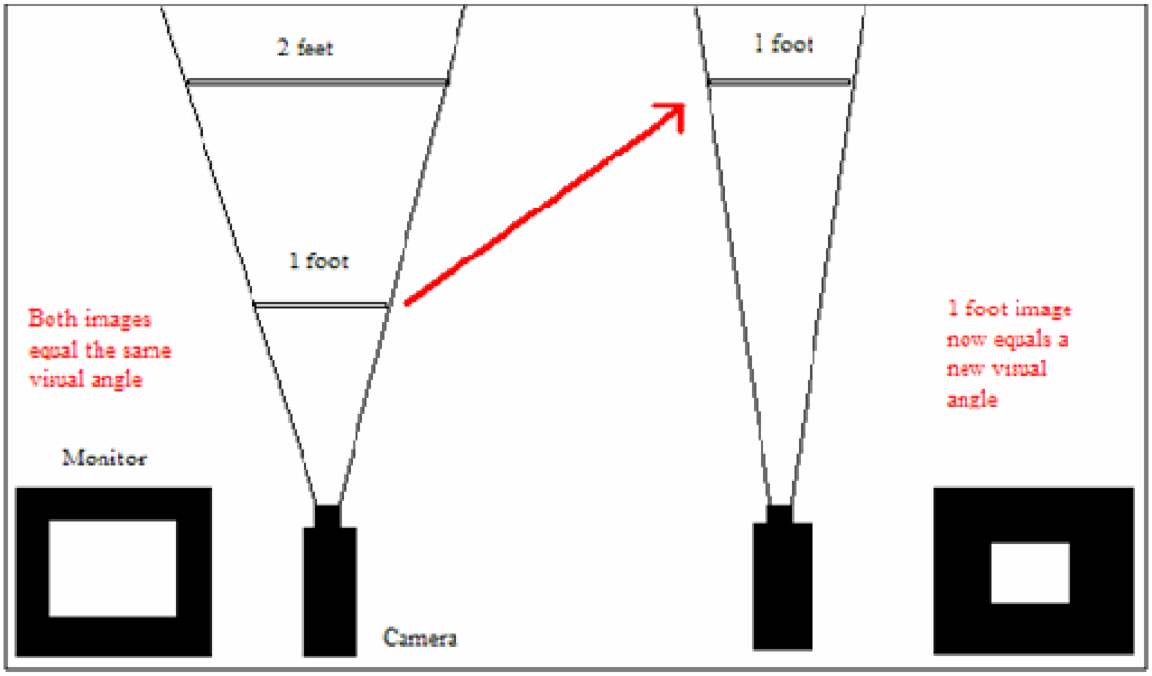

Figure 3: A visual explanation of controlling for visual angle (from Dash, 2004).

In the passive condition, participants used a keyboard to move a remote camera toward and away from the target as they estimated the distance. The camera, mounted on a Puma 560 industrial robotic arm, moved with consistent sinusoidal expansion and contraction. Under the joystick condition, participants completed the same task as in the passive condition, in this case moving the camera arm with a joystick. In the head-coupled condition, participants wore a visor with a lightweight electronic sensor attached and were able to control the movement of the camera by moving their head back and forth.

At the beginning of the experiment the participants took part in a training session in which they viewed three of the target sizes at their respective locations along the reporting device while simultaneously watching them expand and contract on the monitor. This training provided them with an awareness of the expansion and contraction of the image on the screen relating to the distance of the target from the camera lens. After they estimated the distance from the camera lens to the target using a pulley device, they were allowed to view the target again. This feedback on their performance increased their awareness of the radial outflow of the target, based on its actual distance from the camera lens.

Dash hypothesized that head-coupled motion would be superior to the other two conditions of radial outflow. The results, however, indicated no differences between the three conditions, as the slopes of the best-fit lines for simple regressions predicting indicated distance from actual distance being .61, .58, and .59, for the passive, joystick, and head-coupled movement conditions, respectively. While the condition with head-coupled movement did not demonstrate superior accuracy over the passive and joystick conditions, the participants were able to perceive depth from radial outflow under all three conditions. Limitations in technology were cited by Dash (2004) as factors which may explain the lack of an advantage for self-produced movements. Coupling was degraded because of the delay between the head-sensor and the camera movement and interference was also present whenever the participant's head was too close to the monitor in the head-coupled condition. Dash also discussed that the training and feedback, along with the consistent sinusoidal pattern of camera motion is each trial, were critical in allowing the participants to perform in the passive condition. In a separate experiment where no training or feedback was given, the participants were unable to perceive depth in any of the three conditions.

Pagano, et al. (2006) used a similar design to Dash's (2004) involving three passive viewing conditions presented in fixed order of familiar objects with training and feedback, white squares against black space with training and feedback, and white squares with training and no feedback. Simple regressions revealed slopes of .96, .63, and .70, respectively, for the three conditions, indicating that consistent radial outflow produced by an oscillating camera, moving from its center position forward then back then back to center, can provide effective information about depth perception in a remote environment. In addition, their results indicated that after initial training, this information could be effectively used without feedback.

Continuing this line of research, Moore et al. (2006) used the same design for the three conditions of a passive camera, active head movement in time with a passive camera, and head-coupled camera movement. Their results found no differences among conditions, the passive condition being as effective as the active or head-coupled for the perception of distance of targets from a remote camera.

The experiment reported here extends the research of Dash (2004), Pagano et al. (2006), and Moore et al. (2006), by further investigating the applied use of radial outflow as a means of improving perception in remote environments. Participants estimated size and distance perceptions regarding objects viewed via a camera passively moving forward and backward on a robot. Their performance was subsequently compared to a control condition using a static camera. Specifically, during each experimental trial the participants first indicated the size of a target object using a psychophysical device. This was followed by the participants duplicating a specified distance provided by the experimenter by placing the front of the camera lens at that distance from the target objects. The purpose of this experiment was to investigate the following hypotheses:

> **Hypothesis 1:** Participants utilizing a mobile robot with a dynamic camera will produce more accurate size judgments of objects in a remote environment than participants utilizing a mobile robot with a static camera.
>
> **Hypothesis 2:** Participants utilizing a dynamic camera will more accurately position a remote robot when reproducing an indicated distance between the robot and an object than participants utilizing a mobile robot with a static camera.

CHAPTER 2

METHODS

Participants

Two Clemson University graduate students and 10 summer intern undergraduate students participated in this experiment. The ages of participants ranged from 19 to 30, with a mean age of 22. There were six male and six female participants. All participants read and signed an informed consent document prior to participation. All participants were tested for normal visual acuity measured binocularly from 6 m and self-reported full use of their neck, arms, and hands (visual acuity ranged from 20/12.5 to 20/25, with an average of 20/16.8). Participants also reported if they regularly played video games. One participant who was unable to perform the task was replaced. Each participant completed two sessions that were conducted at least 24 hours apart. Each experimental session lasted between 1.5 and 2 hours. Once the experiment was complete, participants were debriefed and received payment for their participation.

Materials

The robot was a modified remote controlled Hobbyzone (Plymouth, MN) model T1A1 Abrams tank (1:16 scale). The vehicle measured 21 cm x 9 cm (see Figure 4).

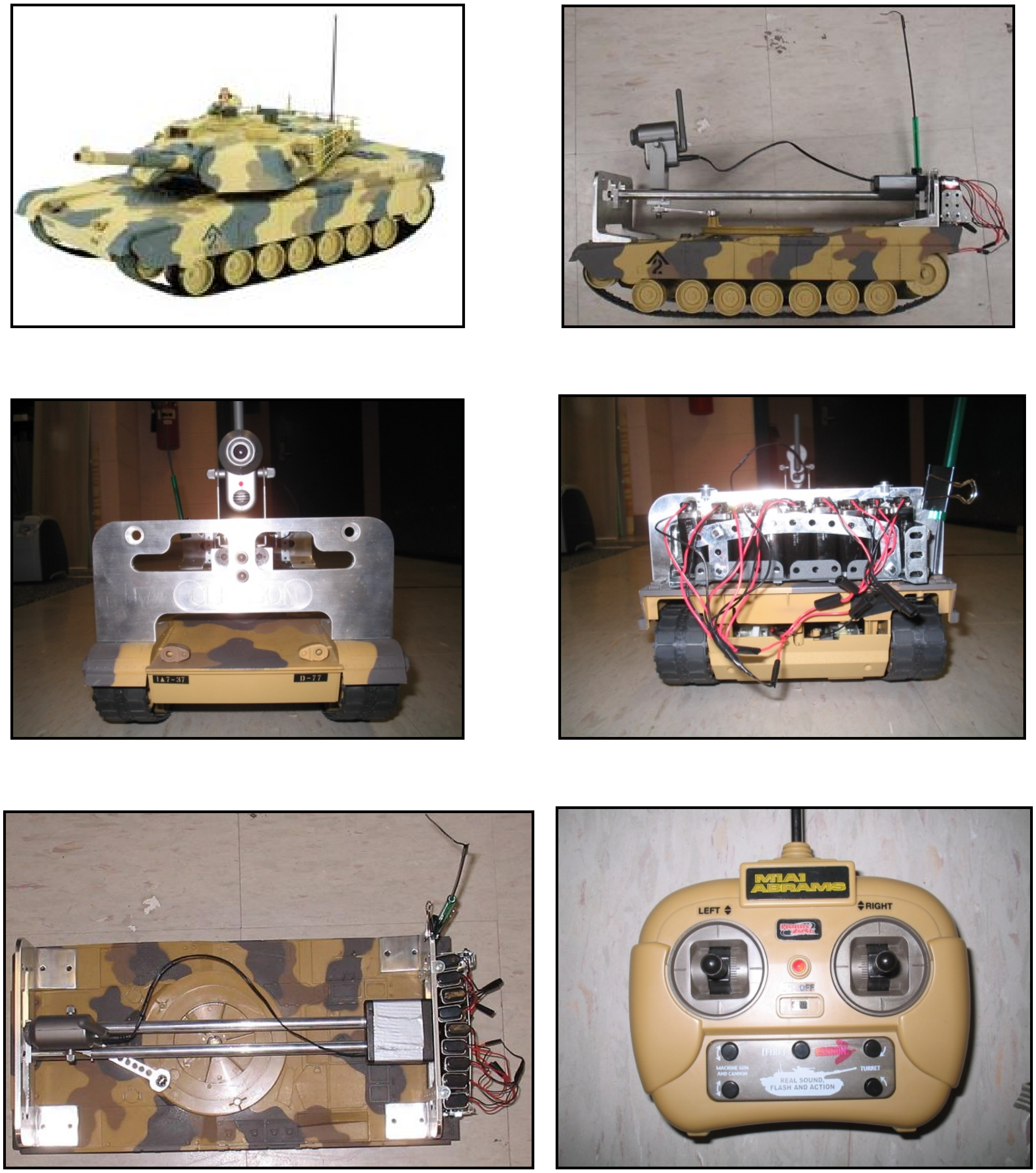

Figure 4: Original and modified RC tank, power supply, camera, & controller.

This vehicle was chosen due to its similarity to tracked robotic vehicles currently used for operations in USAR applications and by the military. This apparatus is an ideal size for laboratory experimentation and is similar to other robots currently used in teleoperation research at Clemson University. The turret on the robotic tank was replaced with a Grandtec USA (Dallas, Texas) wireless “Eye See All” security camera system (see Figure 5). The height of the camera was seven inches from the ground, with an angle of 10 degrees below horizontal.

Figure 5: The Eye See All USB Camera Device.

The camera system included an RF CMOS USB transmitter and receiver. The camera was mounted on rails that allowed it to move forward and backward. This motion was driven by the motor that originally rotated the turret of the tank. These camera motions resulted in generating radial outflow by producing an image that expanded and contracted as the camera moved forward and backward. The camera motions followed a predetermined sinusoidal velocity profile, which was activated by the participant, but not altered by them. The amplitude of movement was 13.5 cm, approximately 64% of the robot's length, with the duration of one oscillation lasting between 3.5 and 4.5 seconds, with higher or lower battery power, respectively. This amplitude is similar to the 17.4 cm voluntary head movements produced by human participants in previous experimentation investigating radial outflow (Pagano & Bingham, 1998). Software for recording the video from the camera was part of the "Eye See All" camera device. Image resolution for the camera device was approximately 320 x 240 pixel array at 30 frames per second with compression. The objects viewed, seen in Figure 6, were 15 white cubes of predetermined sizes, maintaining three different visual angles on the monitor (the specific sizes are given in Appendix A). The objects were placed on a platform against a black background to eliminate visual depth cues. The receiver was installed on a Dell desktop computer and the live camera feed was displayed on a 15" Dell LCD monitor. The resulting image appeared in a 7.25 cm x 9.5 cm window on the monitor for the static camera condition. In the dynamic camera condition, when the camera was in the aft position, the front of the robot came into view. To eliminate this visual cue, foam board was used to reduce the size of the viewing window. The reduced dynamic condition viewing area was 4 cm x 9.5cm as seen in Figure 7.

Figure 6: The smallest and largest target objects.

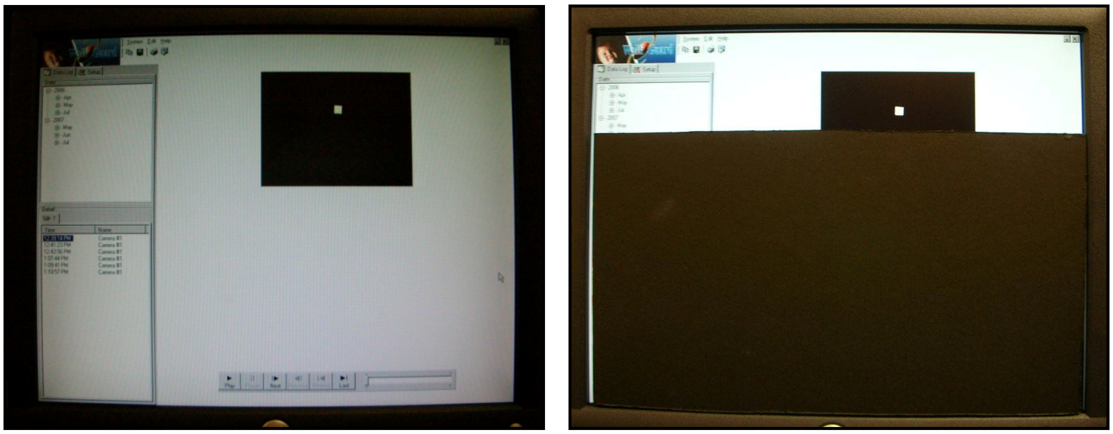

Figure 7: Participant view of practice videos for the static (left) and dynamic (right) conditions.

The remote environment consisted of a 244 cm x 53 cm corridor constructed of flat black wooden walls 61 cm high, with a black foam board floor (see Figure 8). Room lighting, as well as contrast and brightness of the camera view were adjusted so that the targets appeared as white squares against a uniform, black background. The walls and corridor were not visible in the camera view and curtains concealed the remote course and one experimenter from direct participant view.

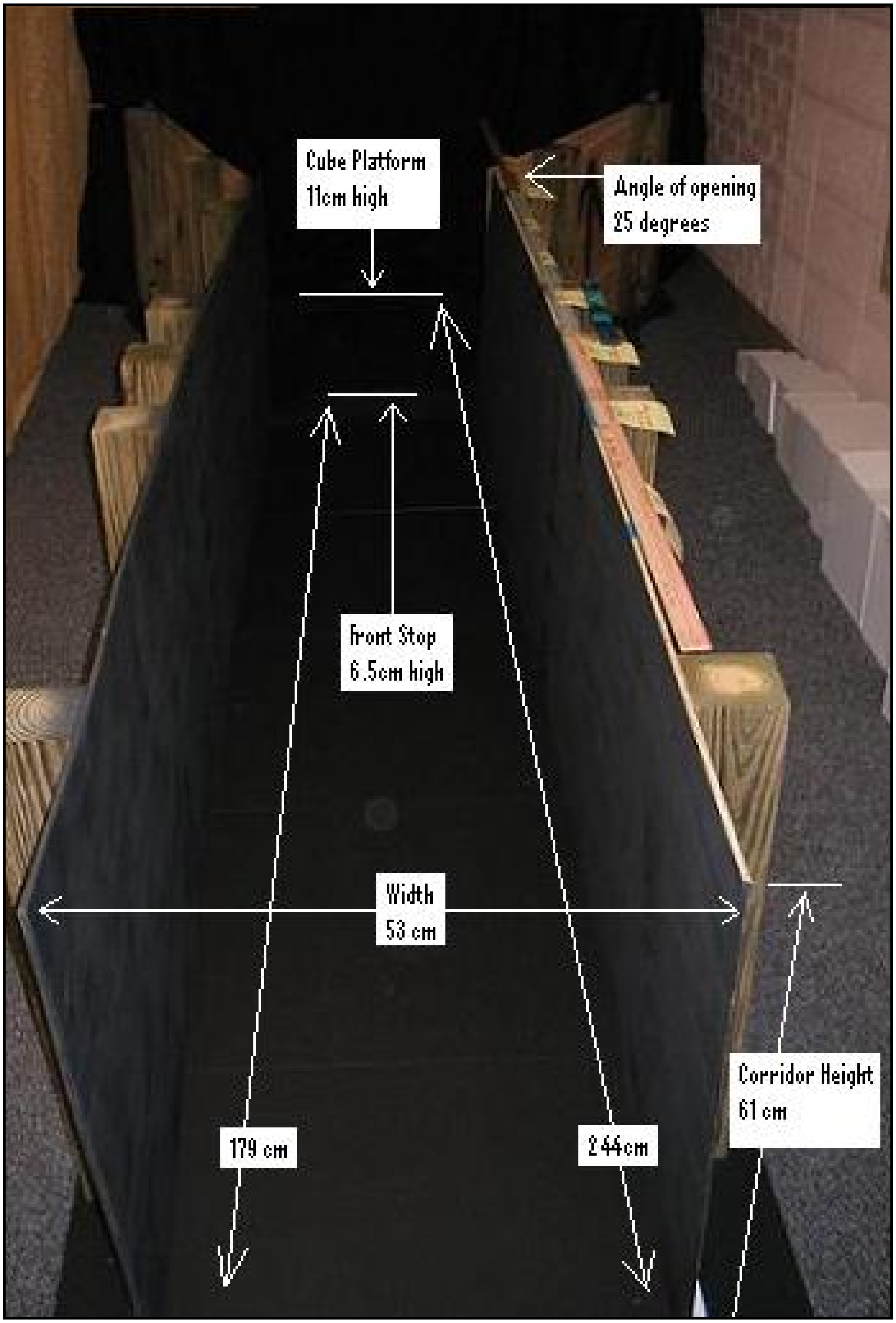

Figure 8: Remote Environment

Participants used a wooden reporting device 4 feet in length for estimating the size of the target objects (see Figure 9). This device was located approximately 30 cm in front of them, parallel to the monitor, with a white foam board background to increase contrast. Following their size estimations participants were instructed to reproduce an indicated distance between the front of the camera lens and the front of the cube. They were instructed to move the tank forward and backward to reproduce the instructed distance. The required distance from tank to target object was provided to the participant by the experimenter via a pulley display device (see Figure 9). The pulley device was perpendicular to the wooden device, to the right of the monitor, allowing participants to view the actual distances next to the view from the robot.

Figure 9: Wooden block and pulley psychophysical reporting devices.

Design

A within-subjects design was used with each participant performing in two separate sessions, one for each camera condition. The two conditions consisted of viewing targets via a static and a dynamic camera. The order for each session was counterbalanced. There were 30 randomized trials for each viewing condition. Dependent measures included size perceptions of the white cubes and distance reproductions of the instructed distances given by the experimenter. Participants reproduced these instructed distances by driving the tank forward and backward.

Procedure

Past research has demonstrated that performance improves with training and feedback (e.g. Bingham et al., 2000; Dash, 2004; Wickelgren et al., 2000), therefore participants were given training time to become familiar with the experiment. Training for participants began with a two to three minute direct-line-of-sight robot familiarization session for forward and backward controls in the static condition, as well as camera movement for the dynamic condition. Training consisted of allowing participants to practice basic robot functionality and view training videos of what the task would look like on the computer screen. These videos illustrated that the size of the target image on the screen was not related to actual cube size. This demonstrated that image size on the screen is decoupled from actual target size and can be misleading.

During the experiment, participants were seated in front of the desktop computer, which displayed the feed from the camera. In front of participants, 27 cm from the edge of the table was the wooden reporting device. This was used by the participant to provide a specific judgment of estimated size of each target object. The participant covered the monitor with black foam board between trials to prevent witnessing the manipulation of the tank distance from the object and the size of the cube being used for each trial.

At the beginning of each trial the robot was placed by the experimenter between 110 and 235 cm from the target. An explanation of starting distances and their derivation is given in Appendix A. The camera was then turned on and the participant drove the robot forward 30 cm at which point the experimenter told them to stop. Participants then judged the size of the target. In the static camera condition, the participants observed the object on the monitor and manipulated the wooden reporting device to give their judgment. In the dynamic camera condition, the participants were instructed to push a button on the remote, which waggled the camera forward and back for two to three oscillations, while watching the object's image expand and contract on the monitor. Following this movement, the participants made their target object size judgments. The camera and the robot never moved simultaneously.

After the size judgment was recorded by the first experimenter, the participant was shown a distance with the pulley display device and was asked to move the robot to reproduce this distance between the target object and the camera lens of the robot. To achieve this, the participant was allowed a maximum of five movements of the tank, either forward or back. In the dynamic camera condition, they were also able to make a maximum of three oscillations of the camera from a stationary position, in between their movements of the tank. After they completed position the robot, they blocked the screen with a piece of black foam board. The second experimenter recorded the reproduced distance, reset the tank and placed the object for the next trial. Feedback was provided to the participant regarding their size judgments and distance reproductions.

CHAPTER 3

RESULTS

Size Perception

Simple regressions predicting perceived target size from actual target size for each of the two viewing conditions are presented in Figure 10. Each data point in Figure 10 represents an average of the two perceived size judgments made by a participant for a given size. The data for all 12 subjects is included in Figure 10. Similar regressions for the individual participants are presented in Tables 1 and 2. The regressions in Table 1 are arranged according to viewing condition, while the regressions in Table 2 are arranged in the order in which they were completed. The regressions in Tables 1 and 2 utilize the data from all trials without averaging. For each participant, perceived size varied as a function of actual size, although the slopes of these functions were far less than 1. Overall, the r^2 values tended to be similar for the two viewing conditions and the slopes of the functions tended to remain consistent over variations in both viewing condition and the order of presentation.

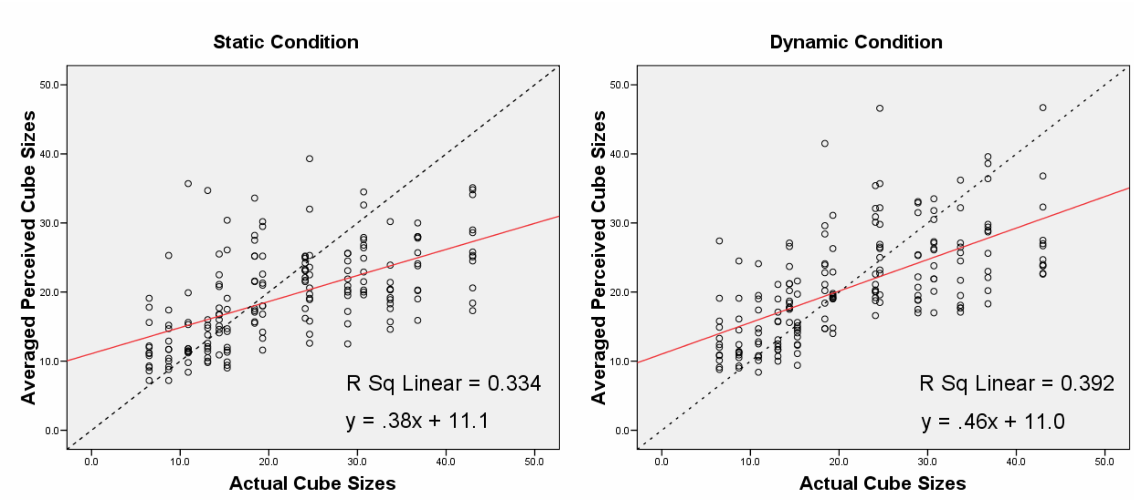

Figure 7: Average perceived cube size as a function of actual cube size for the static and dynamic camera conditions. The red line is the line of best fit for the actual data and the dotted black line is the ideal line of perfect performance with a slope equal to 1 and a Y-intercept equal to 0, in which perceived performance would equal actual performance.

Table 1: Simple regressions predicting perceived size from actual size as a function of camera condition for each of the participants.

Participant	R^2	Slope	Intercept
		Dynamic	
1	.50***	.44	18.3
2	.65***	.30	12.2
3	.38***	.33	11.8
4	.48***	.51	8.3
5	.65***	.64	12.0
6	.65***	.55	7.7
7	.42***	.36	9.4
8	.17*	.30	14.7
9	.38***	.37	8.9
10	.31***	.68	14.4
11	.60***	.55	6.7
12	.54***	.36	9.0
Mean	.48	.45	11.1
All	.31***	.45	11.1
		Static	
1	.25**	.25	17.0
2	.53***	.34	11.0
3	.40***	.24	15.3
4	.63***	.68	6.2
5	.08	-.23	29.2
6	.45***	.36	7.1
7	.59***	.31	8.9
8	.44***	.47	8.2
9	.58***	.64	4.8
10	.41***	.60	9.3
11	.80***	.64	6.4
12	.34***	.24	9.3
Mean	.46	.38	11.1
All	.28***	.38	11.1

$n = 30$ for the individual participants and $n = 360$ overall
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 2: Simple regressions predicting perceived size from actual size as a function of order for each of the participants.

Participant	R^2	Slope	Intercept
	1st Session – Dynamic		
1	.50***	.44	18.3
3	.38***	.33	11.8
5	.65***	.64	12.0
7	.42***	.36	9.4
9	.38***	.37	8.9
11	.60***	.55	6.7
Mean	.49	.45	11.2
Overall	.35***	.45	11.2
	2nd Session – Static		
1	.25**	.25	17.0
3	.40***	.24	15.3
5	.08	-.23	29.3
7	.59***	.31	8.9
9	.58***	.64	4.8
11	.80***	.64	6.4
Mean	.45	.31	13.6
Overall	.21***	.31	13.6
	1nd Session – Static		
2	.53***	.34	11.0
4	.63***	.68	6.2
6	.45***	.36	7.1
8	.44***	.47	8.2
10	.41***	.60	9.3
12	.34***	.24	9.3
Mean	.47	.45	8.5
Overall	.37***	.45	8.5
	2nd Session – Dynamic		
2	.65***	.30	12.2
4	.48***	.51	8.3
6	.65***	.55	7.7
8	.17*	.30	14.7
10	.31***	.68	14.4
12	.54***	.36	8.9
Mean	.47	.45	11.0
Overall	.29***	.45	11.0

n = 30 for the individual participants and n = 180 overall
* $p < .05$, ** $p < .01$, *** $p < .001$

A multiple regression was conducted to determine if the slopes and intercepts differed as a function of camera condition. This multiple regression was performed using actual target size, camera condition (coded orthogonally), and an actual size by camera condition interaction to predict indicated target size. This multiple regression resulted in an $r^2 = .306$ ($n = 720$) with a partial F of 303.2 ($p < .001$) for actual target size, a partial F less than 1 for camera condition ($p > .90$), and a partial F of 2.3 ($p > .10$) for the interaction term. Partial Fs for actual size assess how much the actual target sizes predict the variation in the responses after accounting for the variation due to the other terms. Therefore, the partial F for actual target size tests for a main effect of actual target size. The partial F for condition assesses the degree to which the intercepts for the two sessions differ from each other and test for a main effect of camera condition. The partial F for the interaction term assesses the degree to which the slopes for the two conditions differ from each other. When the multiple regression is repeated without the main effect for camera condition the partial Fs become 303.6 for actual size and 12.8 for the interaction term (both $p < .001$). Although the interaction is significant, it only accounts for approximately 1% of the variance in perceived size, as actual size alone gives an r^2 value of .294. These multiple regressions confirm that the participants perceived size similarly in both camera conditions.

Multiple regressions were also conducted to predict indicated target size from the actual size of the target and the visual angle created by the target when the robot was positioned at the location of the size perceptions. This visual angle quantifies the size of the target's image on the screen from the participant's perspective. The multiple regression resulted in an $r^2 = .369$ ($p < .001$, $n = 720$), with a partial F of 5.0 ($p < .05$) for actual size and a partial F of 84.8 ($p < .001$) for visual angle. This regression was repeated without actual size, resulting in an r^2 of .364, indicating that the participants were basing their size judgments on the visual angle. These multiple regressions give nearly identical results when conducted with the data from the static and dynamic conditions taken individually.

Distance Perception

Simple regressions predicting the reproduced target distances from the instructed target distances for each of the two viewing conditions are presented in Figure 11. Each data point in Figure 11 represents an averaged value of six reproduced distance judgments made by each participant. Thus each data point depicts the judgments for one participant at a particular distance. Similar regressions for individual participants are presented in Tables 3 and 4. The regressions in Table 3 are arranged according to viewing condition, while the regressions in Table 4 are arranged in the order in which they were completed. The regressions in Tables 3 and 4 utilize the data from all trials without averaging. For each participant, reproduced distance varied as a function of instructed distance. Overall, the r^2 values tended to be similar for both viewing conditions and the slopes of the functions tended to remain consistent over variations in viewing condition and the order of presentation.

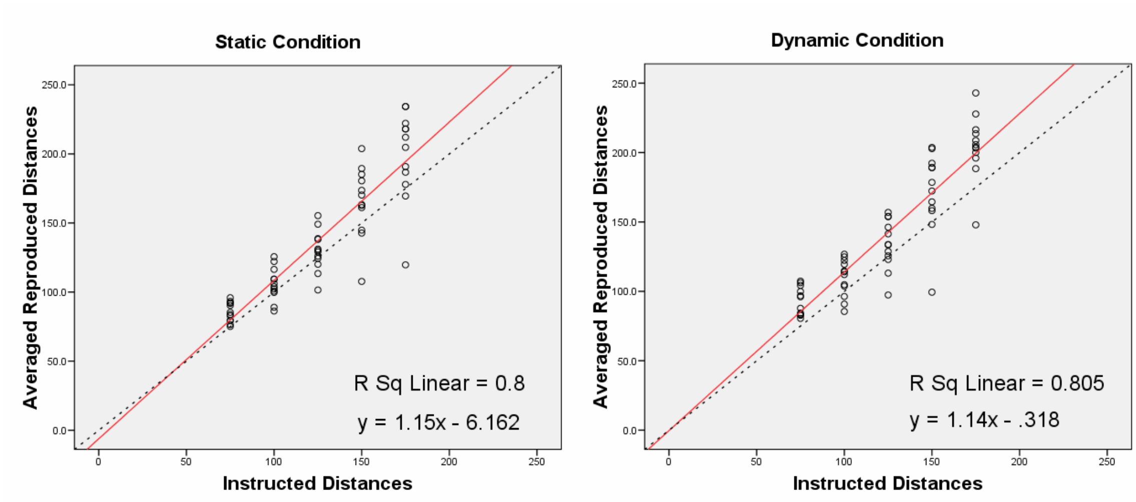

Figure 11: Average reproduced distances as a function of the instructed distances for the static and dynamic camera conditions. The red line is the line of best fit for the actual data and the dotted black line is the ideal line of perfect performance with a slope equal to 1 and a Y-intercept equal to 0, in which perceived performance would equal actual performance.

Table 3: Simple regressions predicting reproduced distance from instructed distance as a function of camera condition for each of the participants.

Participant	R^2	Slope	Intercept (cm)
		Dynamic	
1	.82***	1.24	-4.7
2	.89***	1.20	-11.2
3	.58***	1.09	5.9
4	.82***	1.19	-17.1
5	.88***	1.45	-17.3
6	.87***	1.44	-38.9
7	.72***	1.12	10.8
8	.66***	0.95	22.1
9	.87***	1.37	-9.3
10	.47***	0.60	27.8
11	.68***	0.92	25.2
12	.89***	1.15	2.7
Mean	.76	1.14	-.3
All	.68***	1.14	-.3
		Static	
1	.76***	1.06	-10.1
2	.85***	1.11	-6.5
3	.53***	1.06	9.7
4	.88***	1.45	-38.2
5	.91***	1.62	-45.5
6	.77***	1.34	-30.2
7	.70***	1.30	-25.2
8	.55***	0.98	21.6
9	.83***	1.23	-7.5
10	.44***	0.42	46.3
11	.91***	0.77	32.2
12	.84***	1.42	-20.6
Mean	.75	1.15	-6.2
All	.65***	1.15	-6.2

$n = 30$ for the individual participants and $n = 360$ overall
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4: Simple regressions predicting reproduced distance from instructed distance as a function of order for each of the participants.

Participant	R^2	Slope	Intercept
		1st Session – Dynamic	
1	.82***	1.24	-4.7
3	.58***	1.09	5.9
5	.88***	1.45	-17.3
7	.72***	1.12	10.8
9	.87***	1.37	-9.3
11	.68***	0.92	25.2
Mean	.76	1.20	1.77
Overall	.73***	1.20	1.78
		2nd Session – Static	
1	.76***	1.06	-10.1
3	.53***	1.06	9.7
5	.91***	1.62	-45.5
7	.70***	1.30	-25.2
9	.83***	1.23	-7.5
11	.91***	0.77	32.1
Mean	.77	1.17	-7.75
Overall	.69***	1.17	-7.75
		1st Session – Static	
2	.85***	1.11	-6.5
4	.88***	1.45	-38.2
6	.77***	1.34	-30.2
8	.55***	0.98	21.6
10	.44***	0.42	46.3
12	.84***	1.42	-20.6
Mean	.72	1.12	-4.6
Overall	.61***	1.12	-4.6
		2nd Session – Dynamic	
2	.89***	1.20	-11.1
4	.82***	1.19	-17.1
6	.87***	1.44	-38.9
8	.66***	0.95	22.1
10	.47***	0.60	27.8
12	.89***	1.15	2.7
Mean	.76	1.09	-2.42
Overall	.68***	1.09	-2.43

$n = 30$ for the individual participants and $n = 180$ overall
* $p < .05$, ** $p < .01$, *** $p < .001$

A multiple regression was conducted to determine if the slopes and intercepts differed as a function of camera condition. This multiple regression was performed using instructed distance, camera condition (coded orthogonally), and a reproduced distance by camera condition interaction, to predict reproduced target distance. This multiple regression resulted in an $r^2 = .666$ ($n = 720$) with a partial F (3,717) = 1418.4, $p < .001$, for reproduced target distance and partial Fs less than 1 for camera condition and the interaction term. When the multiple regression is repeated without the interaction term, the partial Fs become 1420.3 ($p < .001$) for instructed distance and 6.2 for camera condition ($p < .05$). Although the main effect of camera condition is significant, it accounts for less than 1% of the variance in reproduced distance, as instructed distance alone gives an r^2 value of .663. These multiple regressions confirm that the participants reproduced distance similarly in both camera conditions.

A multiple regression was also conducted to determine the contributions of instructed distance and size determined by visual angle at the location of the size perceptions. This visual angle provides the size of the object on the screen when the participants were given the instructions to move the robot. This multiple regression resulted in an $r^2 = .675$ ($p < .001$, $n = 720$), with a partial F of 1400.2 ($p < .001$) for instructed distance and a partial F of 28.4 ($p < .001$) for visual angle. The coefficients were 1.13 for instructed distance and 1.74 for the size determined by the visual angle. The simple regression for distance alone resulted in an r^2 of .663, indicating that 66% of the variance in the distances reproduced by the participants can be accounted for by the instructed distances, while only an additional 1% can be accounted for by the visual angle. Similar results are obtained when these regressions are conducted with the data from the static and dynamic conditions taken individually. Thus the distances reproduced by participants can be predicted from the instructed distances alone; they do not seem to be based upon visual angle. These results are somewhat inconsistent with Dash (2004), who found that visual angle accounted for approximately 7% of the variance in distance perceptions.

CHAPTER 4

GENERAL DISCUSSION

The current study was an investigation of size and distance estimations by a teleoperator in a remote environment. Participants operated a robot under two conditions; a static camera and a decoupled, forward and back oscillating camera that provided additional radial flow information. In each session, the operators drove a remote control tank along a course while viewing white cubes. Participants reported size estimations and placed the camera lens of the robot at a designated distance from the target objects.

Overall, participants in this study were unable to determine size accurately under either viewing condition. In addition, size perceptions were not improved by camera motions and were based solely on the size of the visual angle for both conditions. Thus it appears that radial flow produced by the dynamic camera motions did not provide participants with any additional useful perceptual information about the size of an object.

Individual simple regressions for size perceptions of five of the 12 participants' data revealed six slopes that were greater than or equal to .60, with one participant achieving a slope of .64 and an $r^2 = .80$ (see Figure 12 for a depiction of lowest and highest performance). Participant 11 was most accurate at perceiving actual size (see Table 1). Unlike the other participants, participant 11 based their size judgments solely on the actual size of target objects. Thus it may be possible for some participants to determine size under the teleoperation conditions presented in this experiment.

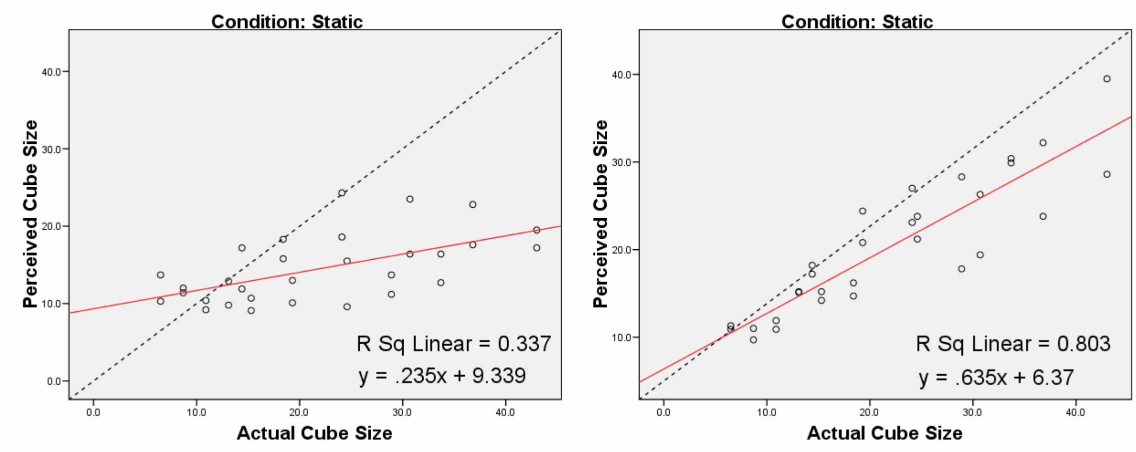

Figure 12: Perceived cube size predicted by actual cube size for the static conditions of the lowest and highest performing participants.

Participants reliably reproduced the instructed robot distances in both viewing conditions. This can be observed by examining the slopes of the regressions presented in Figure 11 and in Tables 3 and 4. Overall, distance reproductions did not change as a function of camera condition and participants did not rely on visual angle in either viewing condition. It can be concluded from this that for distance perceptions, the participants did not utilize the additional information present in the radial flow provided by the dynamic camera. A multiple regression revealed that instructed distance was the primary predictor of reproduced distance. Thus participants had available to them in both viewing conditions information regarding the distance between the robot and the target. The likely source of this information was the radial flow produced by the movements of the entire robot when it was moved forward prior to size judgments and when positioning it for the reproduced distance judgments. One possible explanation of this is that the addition of the camera movements under the dynamic condition did not add any information above what was used by the participants under the static camera condition. Future experiments should investigate limiting the amount of radial flow generated by movement of the entire robot in an attempt to better isolate the effectiveness of a decoupled camera.

The multiple regression for indicated distance resulted in the slope of 1.13 for instructed distance. Thus participants in this study reproduced the instructed distances more accurately than participants in the study by Dash (2004). The slopes obtained in his research were approximately .60. The slope for reproduced distances as a function of instructed distances exceeding 1.0, coupled with an intercept of -16.04, indicates that participants in this experiment were generally overestimating the distance of targets in this experiment and this overestimation generally increased as the instructed distances increased. This increased overestimation with increasing distance is contrary to what was observed by Dash (2004), who found that participants increasingly underestimated target distance as distances increased. Thus it can be concluded that it is possible to accurately determine egocentric distances when teleoperating in a remote environment, if sufficient movement of the vehicle is involved. In Dash's (2004) experiment the dynamic camera was not mounted to a mobile robot. This experiment, however, showed no evidence that a dynamic camera improves determining egocentric distances in a remote environment.

Each of the following should be considered prior to further investigation of radial flow in teleoperation. First of all, in the present experiment the resolution of the camera was only a 76,800 pixel array (320 x 240), which is only 20% of the 380,000 pixels (768 x 494), used by Dash (2004). It is possible that a higher resolution camera would increase the ability of participants to pick up on radial flow provided by the dynamic camera. Second, the sizes of the images presented on the screen used by Dash (2004) were 7.6 cm to 21.3 cm tall. In the present experiment, the sizes were only .5 cm to 2 cm tall. This further reduced the resolution of the information presented to participants. It is possible that if the image sizes were larger, it would increase the ability of participants to pick up on radial outflow information. Third, the range of the camera movement was 4 cm less than that used by Dash (2004). The range of camera motions used by Dash (2004) was 17.4 cm while in this experiment it was 13.4 cm. A larger range of camera motions would have resulted in a greater difference between the sizes of the image at the forward and aft camera positions during the camera motions in the dynamic condition. It is possible that such a larger range of camera motions would have made it easier for participants to pick up on information contained in radial flow.

In addition to the above mentioned shortcomings, there were several other items that added error to the design. First, movement of the camera apparatus was such that it did not exactly follow a consistent sinusoidal velocity profile. Due to the mechanical linkage between the motor and the camera, the apparatus lagged in the aft position slightly and this change in the speed of movement also fluctuated to a small extent, depending on battery power. Second, due to the design, there was a slight rotation of the camera from left to right, of approximately one degree between forward and back movements. Third, while the tank was modified to move slower than its originally designed speed, it was still fast enough to reduce the participants' ability to make fine distance reproductions. This may have contributed noise to the observed distance reproductions.

Individual differences and training are also important factors to consider when examining teleoperation performance. Spatial abilities may play a significant role in teleoperator performance and participants will naturally have differing abilities to control the robotic platform being used. Training is extremely important because robotic platforms do not follow any standardized construction. There are hundreds of robots currently in operation, with varying camera heights, sizes, input devices, capabilities, and constraints. Without specialized training for each of these designs, it would be impossible to navigate or interact remotely and gather a mediated understanding of what is taking place in that environment. Significant effects for the order in which the viewing conditions were presented were not found in this study. However, it is worth noting in Tables 2 and 4 that when participants estimated size in the static condition followed by the dynamic condition, the mean slope and mean r^2 remained constant, while in the opposite order the mean slope dropped by .15 and the mean r^2 decreased slightly. Thus operator performance became more or less accurate despite practice. This was also observed with distance reproductions, as participants who performed in the static condition followed by the dynamic condition actually had a more accurate slope and a larger increase in r^2. This may have implications for training operators who will use robots with a decoupled camera, as decoupled cameras increase the cognitive load of the operator (Hughes et al., 2003). With decoupled cameras it may be necessary to train to proficiency initially with a fixed camera system, prior to training with the same robot and a decoupled camera.

Overall, the most interesting result from this experiment is that participants were able to use radial flow from the movement of the entire robot to perceive both size and distance, despite the limitations of camera resolution and image size of the target objects on the monitor. Furthermore, participants were able to use radial flow provided by a moving robot to perceive distance more accurately than participants in the experiment of Dash (2004), who also used radial flow provided by a dynamic camera. Future research should investigate higher fidelity camera and software systems, increasing the image size of target objects on the screen, and designing situations which may provide advantages of using a decoupled camera in addition to a fixed camera system.

APPENDIX

Three different distances were combined to arrive at controlled trial starting locations for the robot. To generate starting distances between the cubes and the tank camera, distances relating to object size and visual angle were used (Dash, 2004). A randomized distance (5, 10, 15, 20, 25, or 30 cm) was then added to the appropriate distance used by Dash (75, 100, 125, 150, or 175 cm) for each size object. Finally, to provide a baseline amount of optic flow for both conditions, 30 cm was added to the sum of the two numbers above.

Target widths and their respective optical sizes and distances from the initial state of the camera arm in the experiments of Dash (2004). The same widths were used in the generation of cubes in the present experiment.

Target Size (cm^2)	Optical Size (degrees)	Distance (cm)
6.50	5	75
8.70	5	100
10.90	5	125
13.10	5	150
15.30	5	175
14.40	11	75
19.30	11	100
24.10	11	125
28.90	11	150
33.70	11	175
18.42	14	75
24.56	14	100
30.70	14	125
36.83	14	150
42.97	14	175

REFERENCES

Agah, A. & Tanie, K. (1999) Multimedia Human-Computer Interaction for Presence and Exploration in a Telemuseum. *Presence: Teleoperators and Virtual Environments*, *8*, 104-111.

Bingham, G. P. & Pagano, C. C. (1998). The Necessity of a Perception-Action Approach to Definite Distance Perception: Monocular Distance Perception to Guide Reaching. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 145-168.

Bingham, G. P. & Stassen, M. G. (1994). Monocular distance information in optic flow from head movement. *Ecological Psychology*, *6*, 219-238.

Bingham, G.P., Zaal, F., Robin, D., Shull, J.A. (2000). Distortions in Definite Distance and Shape Perception as Measured by Reaching Without and With Haptic Feedback. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1051-1059

Borish, I. (1970). *Clinical refraction* (3rd ed.). Chicago: Professional Press.

Bruckstein, A., Holt, R., Katsman, I., & Rivlin, E. (2005). Head Movements for Depth Perception: Praying Mantis vs. Pigeon. *Autonomous Robots, 18*, 21-42.

Burke, J., Murphy, R., Coovert, M., & Riddle, D. (2004). Moonlight in Miami: A Field Study of Human-Robot Interaction in the Context of an Urban Search and Rescue Disaster Response Training Exercise. *Human-Computer Interaction*, *19*, 85-116.

Casper, J. & Murphy, R.R. (2003) Human-Robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, *33*, 367-385.

Dash, C. Self-produced motion as an effective depth cue for teleoperation. Unpublished Masters thesis, Clemson University. August, 2004.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Houghton Mifflin, Boston.

Goldstein, E.B. (2007). *Sensation and Perception*. 7th Edition. Thomson, Belmont, CA.

Hughes, S., Manojlovich, J., Lewis, M., and Gennari, J. (2003). Camera Control and Decoupled Motion for Teleoperation. *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics, Washington, DC,* 2, 1339-1344.

Lee, D. N. (1974). Visual information during locomotion. In R. B. McLeod & H. L. Pick (Eds.), *Studies in perception: Essays in honor of J. J. Gibson* (pp. 250-267). Ithaca, NY: Cornell University Press.

Lin, Q. & Kuo, C. (1999). Assisting the teleoperation of an unmanned underwater vehicle using a synthetic subsea scenario. *Presence: Teleoperators & Virtual Environments, 8*, 520-530.

Mailhot, C.B. (1996). Telepresence: What is it? *Nursing Management*. *27*, 32I,32L.

Moore, K., Gomer, J., Rodes, W., Chitrakaran, V., & Pagano, C. (2006). Using Radial Outflow to Enhance Depth Perception During Remote Viewing. Poster presented at the North American Meeting of the International Society for Ecological Psychology, Cincinnati, OH.

Murphy, R. (1999). Case Studies of Applying Gibson's Ecological Approach to Mobile Robots. IEEE Transactions on systems, man, and cybernetics – Part A: Systems and Humans, *29*, 105-111.

Murphy, R. (2004). Trial by Fire: Activities of the Rescue Robots at the World Trade Center from 11-21 September 2001. *IEEE Robotics & Automation Magazine*. September 2004, 50-60.

Negahdaripour, S. & Madjidi, H. (2003). Stereovision Imaging on Submersible Platforms for 3-D Mapping of Benthic Habitats and Sea-floor Structures. *IEEE Journal of Oceanic Engineering 28*, 625-650.

Pagano, C.C. & Bingham (1998) Comparing Measures of Monocular Distance Perception: Verbal and Reaching Errors are not Correlated. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1037 – 1051.

Pagano, C., Smart, M., Blanding, T., Chitrakaran, V. (2006). The Use of Radial Outflow for the Perception of Depth in Remote Environments. Poster presented at the meeting of the Vision Sciences Society, Sarasota, FL.

Reinhardt-Rutland, A.H. (1996). Remote operation: A selective review of research into visual depth perception. *The Journal of General Psychology, 123*, 237-248.

Rogers, B. & Graham, M. (1979) Motion parallax as an independent cue for depth perception. *Perception, 8*, 202-214.

Sekman, A., Wilkes, M., Goldman, S., & Zein-Sabatto, S. (2003). Exploring importance of location and prior knowledge of environment on mobile robot control. *International Journal of Human-Computer Studies*. *58*, 5-20.

Sheridan, T.B. (1989). Telerobotics. *Automatica, 25*, 487-507.

Sheridan, T.B. (1992). *Telerobotics, Automation and Human Supervisory Control.* MIT Press, Cambridge, M.A., pg. 393.

Smets, G., Overbeeke, C., & Stratmann, M. (1987) Depth on a Flat Screen. *Perceptual and Motor Skills, 64*, 1023-1034.

Smets, G. (1995). Designing for Telepresence: The Delft Virtual Window System. *Local Applications of the Ecological Approach to Human-Machine Systems*, *2*, 182-207.

Stassen, H.G. & Smets, G.J.F. (1997). Teleoperation and telepresence. *Control Eng. Practice, 5*, 363-374.

Tittle, J.S., Roesler, A. & Woods, D.D. (2002). The remote perception problem. *Proceedings of the Human Factors and Ergonomics society 46th annual meeting, Baltimore, MD*, 260-264.

Wickelgren, E.A., Daniel, M.C., Bingham, G.P (2000). Reaching measures of monocular distance perception: Forward versus side-to-side head movements and haptic feedback. *Perception and Psychophysics. 62*, 1051-1059.

Woods, D.D., Tittle, J., Feil, M, & Roesler, A. (2004). Envisioning human-robot coordination in future operations. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, *34*, 210-218.