2012

Multimodal Human-Machine Interface For Haptic-Controlled Excavators

Benjamin Osafo-Yeboah
North Carolina Agricultural and Technical State University

MULTIMODAL HUMAN-MACHINE INTERFACE FOR HAPTIC-CONTROLLED EXCAVATORS

by

Benjamin Osafo-Yeboah

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Department: Industrial and Systems Engineering
Major: Industrial and Systems Engineering
Major Professor: Dr. Steven Jiang

North Carolina A&T State University
Greensboro, North Carolina
2012

ABSTRACT

Osafo-Yeboah, Benjamin. MULTIMODAL HUMAN-MACHINE INTERFACE FOR HAPTIC-CONTROLLED EXCAVATORS. (**Major Professor: Steven Jiang**), North Carolina Agricultural and Technical State University.

Since the 1940s, fluid power has been used effectively in combination with other technologies to provide power in the form of hydraulics or pneumatics for a variety of industries. One such machinery is the excavator. Although the excavator is widely used in industry, numerous design constraints make the interface less intuitive, resulting in long operator training and high cost. Further, traditional excavator-operator interfaces rely mainly on visual and to some extent auditory senses. This often leads to cognitive overload with its negative effect on performance. A haptic-controlled excavator interface has been proposed as an alternative to the traditional excavator interface.

The goal of this research is to develop a human-excavator interface for the haptic-controlled excavator that makes use of the multiple human sensing modalities (visual, auditory haptic), and efficiently integrates these modalities to ensure intuitive, efficient interface that is easy to learn and use, and is responsive to operator commands. Two empirical studies were conducted to investigate conflict in the haptic-controlled excavator interface and identify the level of force feedback for best operator performance. A quantitative model of human interaction with haptic-controlled excavator was developed. Design recommendations to improve the existing haptic-controlled excavator interface were identified using interface design guidelines. Finally, an evaluation of the modified haptic-controlled excavator interface was conducted to assess operator performance and to identify potential usability problems.

School of Graduate Studies
North Carolina Agricultural and Technical State University

This is to certify that the Doctoral Dissertation of

Benjamin Osafo-Yeboah

has met the dissertation requirements of
North Carolina Agricultural and Technical State University

Greensboro, North Carolina
2012

Approved by:

Dr. Steven Jiang Major Professor	Dr. Celestine Ntuen Committee Member
Dr. Silvanus Udoka Committee Member	Dr. Zongliang Jiang Committee Member
Dr. Paul Stanfield Department Chairperson	Dr. Sanjiv Sarin Associate Vice Chancellor for Research and Graduate Dean

BIOGRAPHICAL SKETCH

Benjamin Osafo-Yeboah, a Sloan Scholar, was born on sunny afternoon of September 17, 1973 in the beautiful city of Accra, Ghana to Francis Yeboah and Helena Pomaah. He obtained his General Certificate of Examination (GCE) Advanced Level Certificate from Saint Peter's Secondary School, and enrolled at the University of Science and Technology (UST) in Kumasi, Ghana, where he received a bachelors' degree in Engineering Geology in 1998.

In May 2002, Benjamin moved to the United States and worked with Limited Brands Logistics Inc. in Columbus, Ohio. In August 2005, he enrolled into the graduate program in Industrial & Systems Engineering at North Carolina Agricultural and Technical State University (NCA&T) and obtained a master's degree in Industrial & Systems Engineering in December, 2007, with specialty in Human Factors. In January 2008, Benjamin enrolled into the doctoral program in Industrial & Systems Engineering at NCA&T and is due to receive his doctoral degree in Industrial & Systems Engineering (Human Factors Specialty) in May 2012. Benjamin has conducted research and published several peer-reviewed journal articles and conference proceedings in the areas of human-machine interaction, usability and human performance modeling. He has acquired useful industry experience working as an Industrial Engineer for Boston Scientific in the summer of 2010 and 2011. He has received several honors and awards including the most outstanding graduate research assistant award in April, 2009 and winner of graduate students' technical paper competition at HICS conference in April, 2008, in Norfolk, VA.

ACKNOWLEDGEMENTS

I would like to thank everyone who has supported, encouraged and counseled me during this process. In particular, I would like to extend my sincere gratitude to my advisor, Dr. Steven Jiang, for his support, patience and guidance throughout my stay here at North Carolina A&T State University (NCA&T). Thank you for your mentorship and friendship. I would like to thank my dissertation committee, Dr. Celestine Ntuen, Dr. Silvanus Udoka and Dr. Zongliang Jiang for their support and guidance.

I would also like to thank my colleagues Ritson Delpish, Hanniebey Wiyor, Dr. Jules Chenou and all my peers at the Center for Human-Machine Systems Laboratory for their support and encouragement. You encouraged and brainstormed with me when I needed help most. I will forever be grateful.

I will also like to say thank you to the Center for Compact and Efficient Fluid Power Systems (CCEFP) teams at NCAT and at the Georgia Institute of Technology (Georgia Tech). In particular, I would like to thank Mark Elton of Georgia Tech who developed the *Matlab* program that run the haptic-controlled excavator simulation used in this research. Mark also provided all the necessary help that enabled me to conduct usability testing of the haptic-controlled excavator interface at Georgia Tech. Thank you so much Mark. Lastly, I would like to thank CCEFP for funding this research.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE **PAGE**

LIST OF TABLES

TABLE **PAGE**

CHAPTER 1

INTRODUCTION

1.1 Introduction

The term multimodal interface has become a buzzword used in many contexts and across several disciplines (Bernsen, 1994, 1997). Multimodal human–machine interaction lies at the crossroads of several research areas including computer science, psychology, artificial intelligence, human factors, ergonomics etc. Multimodal human-machine interface studies allow researchers to determine how computer technology could be made more usable by people, which requires the understanding of at least three things: the user who interacts with it, the system (the computer technology and its usability), and the interaction between the user and the system (Jaimes & Sebe, 2007). As Oviatt (1999) puts it, multimodal systems represent a research-level paradigm shift away from conventional windows-icons-menus- pointers (WIMP) interfaces toward providing users with greater expressive power, naturalness, flexibility, and portability. Well-designed multimodal systems integrate complementary modalities to yield a highly synergistic blend in which the strengths of each mode are capitalized upon and used to overcome weaknesses in the other (Oviatt, 1999).

Traditionally, human-machine interfaces rely exclusively on visual modality (e.g. keyboards, displays, levers, pedals) and auditory modality (e.g. alarms) as the pathway for communication between humans and machines. However, as operator workload and task are increased, the visual modality becomes overloaded due to the limited number of channels through which the machine and the operator can communicate. The over-working of the visual and speech modalities lead to operator fatigue which ultimately results in under-performance and errors.

A multimodal human-machine interface aims to develop the necessary technology that will provide a more intuitive and natural way for people to operate and control computers and machines. It allows users to control and interact with machines using multiple input modalities including speech, sight, touch, taste, smell, gestures. A multimodal human-machine interface therefore, has the potential to minimize the user's cognitive workload when performing complex tasks as attentional resources will be drawn from different resource pools.

Recent research in human-machine interface has concentrated on investigating the cognitive behavior of the operator in order to find ways to reduce his/her mental workload. The use of multiple modalities for human-computer interaction will improve the nature of human-computer collaboration as the computer in human-computer interaction system becomes more of an active participant in the task (Schomaker et al., 1995). To achieve these goals, appropriate interfaces that incorporate all or most sensing modalities such as visual, tactile, auditory, taste will be necessary. For example, Schomaker et al. (1995) suggests that if the task at hand involves the manipulation of objects, then an appropriate interface is likely to use a combination of visual, auditory and haptic/tactile information.

Advances in technology have made multimodal interactive systems more feasible unlike the 1960s and 1970s when manual control modeling was the order of the day. This has led to the development of reasonable and user-friendly interfaces for various interactive applications resembling natural human-computer interaction. For example, in natural human-human interaction, verbal information exchanged alongside non-verbal signs, gestures, facial expressions and other cues that compliment the information exchange. As technology advances and programming languages becomes more intuitive, it is hoped that more of the cues and gestures that make human-human conversations so effective could be incorporated into programming language to further enhance the next generation human-computer interaction.

While it is envisioned that multimodal human machine interfaces would have a profound impact on how humans interact with machines in the future, there are several potential troubling issues and disadvantages that need to be addressed in order to realize the full potential of multimodal human-machine interfaces. Among these problems is the issue of coordination and combination of multiple modalities. Hurtig & Jokinen (2006) suggest that special attention must be paid to the system on interpretation level and from the point of view of usability, since there is a danger that the users might be exposed to cognitive overload by the stimulation of too many media. For example, in route navigation tasks, they suggest that the system should guide users accurately and quickly and provide necessary assistance in tasks that are complicated and confusing. A study to investigate operator behavior while using a haptic-controlled excavator simulator by Osafo-Yeboah et al. (2009) found that there was great interdependence among certain operator behavior events, suggesting a possible struggle to coordinate operator's hand and eye movements. Issues of coordination thus need to be properly investigated when designing multimodal human machine interface to enable users realize their full potential.

Oviatt (1999), however, cautions that multiple modalities alone do not bring benefits to the interface and may be ineffective or even disadvantageous.

The five human senses are sight, touch, hearing, smell, and taste, and the input modalities of most computer input devices correspond to these senses. For example, cameras use the sense of sight, haptic sensors use the sense of touch, microphones use the sense of hearing, olfactory uses the sense of smell, and flavor uses the sense of taste. However, there are many other computer input devices activated by humans that do not fall under any of the categories described above but fall under a combination of the human senses, for example, keyboard, mouse, writing tablet, biometric sensors etc. (Legin et al., 2005). Designing an efficient and user-friendly multimodal human-machine interface requires experts with diverse backgrounds in topics such as psychology and cognitive science to understand the user's perceptual, cognitive, and problem solving skills; sociology to understand the wider context of interaction; ergonomics to understand the user's physical capabilities; graphic design to produce effective interface presentation; and computer science and engineering to be able to build the necessary technology (Jaimes & Sebe, 2007).

1.2 Multimodal Application to Fluid Power Systems

Since the 1940s, fluid power has been used effectively in combination with other technologies through the use of sensors, transducers and microprocessors to provide power in the form of hydraulics and/or pneumatics for a variety of industries. Since their introduction, fluid power systems have advanced progressively to meet the constant demand for new technology that accomplishes tasks more easily, efficiently, and economically. Hydraulic systems are important actuators in modern industry, principally because they have a high power/mass ratio, fast response, and high stiffness: a combination unmatched by any other commercial technology (Alleyne & Liu, 2000). Industries that have benefited the most from advances in fluid power technology include agriculture, construction, manufacturing, mining, transportation, and aerospace.

One particular application of fluid power technology that has seen major improvements in terms of engine performance, better operator interface and ergonomically safe cab design is the excavator (Carter, 2008). The excavator has numerous applications in the construction, mining, agricultural and transportation industries. The excavator is an earthmoving equipment that is powered by hydraulics, and consists of digging bucket attached to the end of a movable, articulated arm that can be used to tackle a wide variety of trenching, loading, scooping, filling, and leveling chores that would otherwise require multiple machines and considerably more time. An example of excavator is seen in Figure 1.1.

Like most other earthmoving machinery, operating an excavator is not an easy task. First problem is the need for operators to solve the inverse kinematic relationships between lever displacement and bucket trajectory (Frankel, 2004). Secondly, the dual-ended nature of excavators present concerns about operator ergonomics, visibility as well as comfort. A good design must ensure that operators have unrestricted sightlines, perform tasks comfortably and be in control of the equipment no matter which end of the machine they are operating. Further, fluid power systems present other issues such as high pressure, friction, containment and constant movement, which lead to problems in controllability, leakage and efficiency. These problems are compounded by the fact that excavator operation requires direct manual control, which in turn requires excessive amounts of energy, intense task concentration, and high skill level to accomplish.

Figure 1.1: Bobcat backhoe excavator

Due to these constraints, excavator operators have to be trained for long periods of time before they are able to comfortably operate the machine and solve the inverse kinematic relationships subconsciously. This requires great deal of skills, concentration, and effort on the part of the operator to accomplish. Since the only feedback available to the operator is the observed bucket speed, the engine's response to a load, and/or pressure waves propagated back to the user's hand, it is usually not easy for novice operators to have a 'feel' for the non-intuitive level motions (Kontz & Book, 2007). As a result, construction companies often have to hire or contract professional operators for even the simplest earthmoving tasks usually at a high cost and inconvenience.

To overcome this problem, the haptic interface is being considered as an alternative to the traditional direct manual manipulating control levers. Since human cognitive processes and perception build largely upon multimodality, a proper combination of different interface components will result in a flow of information on several parallel channels and has been shown to enhance effectiveness of interaction (Krapichler et al., 1999). By making use of the haptic control interface instead of the traditional levers and pedals, excavator operators will be freed from solving the inverse kinematic relationship and therefore, help them perform their tasks more effectively, and also shorten the training time for novice operators (Kontz & Book, 2007).

By using special input/output devices (joysticks, stylus or other devices), operators can receive feedback in the form of 'feel' sensation in the hand while operating the excavator. In combination with visual display, haptic interface can be used to train operators to better perform digging tasks requiring hand-eye coordination, and provide valuable help to novice operators with little experience to improve their task performance. The excavator, like most traditional earthmoving equipment relies on visual and/or auditory feedback, providing operators with only two modalities. By incorporating the haptic interface into the new design, a third modality, 'haptic feedback' is introduced into the design with the expectation that this extra modality will compliment the other two modalities (visual and auditory) and result in improved operator performance. Also, haptic feedback will help excavator operators to avoid damaging utility lines during excavation, by providing a force feedback that alerts the operator to the presence of unusual obstacles whenever such utilities lines are encountered and, therefore, lead to a safer use (Osafo-Yeboah et al., 2009). Further, such a multimodal designed excavator interface will be intuitive, easy to use and can reduce operator mental workload and stress level leading to improved situation awareness, facilitate depth judgment, and speed up decision making resulting in improved performance.

1.3 Outline of Research Problem

1.3.1 Research Goals.

This research seeks to investigate how audio, visual and tactile sensory modalities could be incorporated into the design of a haptic-controlled excavator interface. A haptic-controlled excavator interface testbed currently under development as part of a broader effort to develop efficient, effective and safe fluid power systems will be used to model, characterize and experiment on multimodal human-machine interfaces for emerging fluid power actuated devices, by taking advantage of the multiple sensing and display modalities to enhance operational effectiveness. To achieve this, two empirical studies will be conducted to investigate first, the existence of conflict and interference in a multimodal haptic-controlled excavator interface, second, to investigate whether force feedback improves operator performance when using haptic-controlled excavator, and third, to identify the range of force feedback values necessary for best operator performance. In these empirical studies, performance measures will be task completion time, number of scoops required to fill bin, and rate of accuracy, A quantitative model will be developed to predict operator performance in tasks involving haptic-controlled excavator interface. Results from the two empirical studies together with interface design principles will be recommended for use in the design and development of a more intuitive, safe, efficient and effective interface for excavators and other fluid power systems. The following sub-goals are outlined to help achieve the primary objective.

1. Conduct an empirical research to investigate whether conflict exists between the haptic, visual and auditory modalities in a multimodal haptic-controlled excavator interface, and if conflicts do exist, how their impact could be minimized in the domain of haptic-controlled excavator interface.
2. Conduct an empirical research to investigate the impact of haptic force feedback on operator performance in the haptic-controlled excavator interface domain, and to identify the level/range of force feedback values that result in preferred operator performance.
3. Develop and implement a quantitative model that allows the prediction of operator performance in tasks involving haptic-controlled excavator interface through control theoretic approach. Develop a proof-of-concept based haptic-controlled excavator interface model, implement model in Matlab and validate quantitative model results by comparing to experimental results.
4. Use results from empirical studies, interface design principles and usability guidelines to recommend improvements to multimodal haptic-controlled excavator interface for safe, efficient, effective user interaction.

Specifically, the following research questions will be answered: (1) Does conflict exist between sensory modalities (auditory, visual, and haptic) in the haptic-controlled excavator interface? (2) Are these conflicts significant enough to have an impact on the performance of operator-excavator interaction? (3) Do operators have problems coordinating their hand-eye movement? If they do, how would this affect the efficient operation of the haptic-controlled excavator? (4) Does different force feedback affect operator performance? (5) What is the range of force feedback values that yield optimal operator performance?

1.3.2 Intellectual Merit.

This dissertation research will yield the following tangible contributions:

1. This work will provide a comprehensive literature review of multimodal human machine interface systems, and the important role of the five human senses in designing effective, efficient, safe and intuitive user interfaces
2. This work will provide a comprehensive understanding of whether or not conflicts do exist between visual, haptic and auditory modalities in the domain of multimodal human-excavator interface (haptic-controlled excavator interface), and how these conflicts affect operator performance. Recommendations from this work will provide valuable information to engineers and designers as they attempt to develop a truly intuitive, safe and efficient multimodal interfaces for excavators and other emerging fluid power systems.
3. This work will provide empirical evidence to support the claim that the use of force feedback in a haptic-controlled excavator interface impacts operator performance. Further, this work will provide empirical evidence to identify the range of force feedback values that produces the optimal operator performance in a haptic-controlled excavator interface domain.

4. While there are control theory models that predict operator performance in pursuit tracking tasks such as in a piloting task, there are no such models to predict operator performance in a multimodal haptic-controlled excavator interface domain. This work developed and implemented a haptic-controlled human excavator model to predict operator performance in a multimodal haptic-controlled excavator domain.
5. This work developed a multimodal human-machine interface design framework for excavators and other fluid power systems of the future, and performed usability evaluation of the improved haptic-controlled excavator interface.

1.4 Chapter Summary

Multimodal human-machine interfaces present opportunity for engineers and designers to develop interfaces/products that are intuitive and thus allow users to have safe, efficient, natural, and fulfilling interaction with the system. This research provides an overview of multimodal human-machine interfaces and its application to fluid power systems. Specifically, a multimodal haptic-controlled human-excavator interface model, currently under construction at the Georgia Institute of Technology by the Center for Compact and Efficient Fluid Power Systems (CCEFP) is investigated as part of a larger effort to develop an efficient, safe and effective alternative to the traditional pedals and levers excavator interface.

Empirical investigations are conducted to: (1) assess conflict and interference between visual, haptic and auditory modalities in the haptic-controlled excavator interface, and their impact on operator performance, (2) determine whether force feedback in the haptic-controlled excavator interface affects performance, (3) identify the range of force feedback values that produce the best operator performance.

Results from the empirical studies are used to modify and improve the current haptic-controlled excavator interface, and usability evaluation is conducted to assess the performance of the improved interface. Further, a quantitative model is developed using control theory approach to help predict operator performance in excavation task while using the haptic-controlled excavator interface.

1.5 Organization of Dissertation

The rest of this dissertation is organized as follows. Chapter 2 reviews multimodal human-machine interfaces and their basic theories. It compares these theories and summarizes their design challenges. Chapter 3 surveys conflicts in multimodal human-machines interfaces, as well as an empirical investigation of conflicts in a haptic-controlled excavator interface. In Chapter 4, an empirical study is conducted to investigate the impact of force feedback on performance in a haptic-controlled excavator interface. Chapter 5 presents a brief description of control theory and fuzzy logic model of perception, and presents a quantitative model to predict performance in a haptic-controlled excavator interface. Chapter 6 presents an interface design framework for multimodal interfaces and combines recommendations from empirical studies to make design changes to existing haptic-controlled excavator interface. Finally, Chapter 7 presents general discussion and conclusions as well as the rationale and the major contributions of the research.

CHAPTER 2

LITERATURE REVIEW

There are five human sensory modalities: sight, audio, tactile, taste and smell, however, it should be noted that at the neurophysiologic level, there are several identifiable input channels through which humans perceive. A summary of the simplified version of the input channels at the neurophysiologic level (Schomaker et al., 1995) is shown in Table 2.1. Though these input channels/modalities exist at the neurophysiologic level, not all are of interest when it comes to human computer interaction mainly because they do not have cortical representation (sense of balance and chemical senses) or they may have a very reduced one (taste modality) and do not give origin to conscious perception (Schomaker et al., 1995).

Though taste and smell are described here as two of the five human senses, they are not very useful channels for human computer interaction due to the impractical nature of their applications in human computer interaction settings. Multimodal human-machine interface provides the users with multiple modalities with which they can interact with a system beyond the traditional keyboard and mouse input/output. A well-designed multimodal system integrates complementary modalities to yield a highly synergistic blend in which the strengths of each modality is capitalized upon and used to overcome weaknesses in the other (Oviatt, 1999) . Other advantages of using multiple modalities include increased usability, error prevention, robustness of the interface, helping users recover from errors, bringing more bandwidth to the communication, and adding alternative communication methods to different situations and environments (Jaimes &

Sebe, 2007). However, in practice, the modalities of seeing and hearing are the most commonly employed.

Table 2.1: Overview of input channels at the neurophysiologic level
[Courtesy of Schomaker et al., 1995]

Sensory Modality	**Form of Energy**	**Receptor organ**	**Receptor Cell**
Chemical (Internal)			
blood oxygen glucose pH (cerebrospinal fluid)	O_2 tension carbohydrate oxidation ions	carotid body hypothalamus medulla	nerve endings gluco-receptors ventricle cells
Chemical (external)			
taste smell	ions and molecules molecules	tongue and pharynx nose	Taste bud cells olfactory receptors
Somatic Senses			
touch pressure temperature pain	mechanical mechanical thermal various	skin skin and deep tissue skin, hypothalamus skin & various organs	nerve terminal encapsulated nerve endings peripheral & central nerve terminal
Muscle sense, kinesthesia			
muscle stretch muscle tension joint position	mechanical mechanical mechanical	muscle spindles tendon organs joint capsule & ligaments	nerve terminal nerve terminal nerve terminal
Sense of balance			
linear acceleration angular acceleration	mechanical mechanical	sacculus/utriculus sacculus/utriculus	hair cells hair cells
Hearing			
	mechanical	cochlea	hair cells
Vision			
	light	retina	photoreceptors

In recent years, there has been a great deal of research in the area of haptic feedback and control, though taste and smell sensing modalities have not received as much attention. The increased interest generated in haptic input devices has opened up new possibilities for the fluid power industry. It is hoped that this will result in more intuitive designs of fluid power equipment which will enable even novice users to become proficient at operating heavy earthmoving equipment more quickly and efficiently than was previously possible (Frankel, 2004). The following sections describe each of the five human sensing modalities in detail.

2.1 Visual Modality

Vision is the physiological sense of sight by which the form, color, size, movements, and distance of objects are perceived. In other words, vision is the ability to see. Webster's dictionary describes vision as 'the special sense by which the qualities of an object constituting its appearance are perceived and which is mediated by the eye'. The visual system in humans allows individuals to absorb information from the environment. It is part of the central nervous system which enables humans to process detail visual information as well as several non-image forming photo response functions. The visual system interprets information from visible light to build a representation of the surrounding world. It also accomplishes a number of complex and non-image forming tasks, such as the reception of light and the formation of monocular representations; the construction of a binocular perception from a pair of two dimensional projections; the identification and categorization of visual objects; assessing distances to and between objects; and guiding body movements in relation to visual objects.

As an input modality for information processing, vision plays the most important role, and is traditionally believed to dominate haptic and auditory senses in object perception, however, recent studies have shown that object perception is a lot more complicated and depends on the situation (Locher, 1982; Sathian & Zangaladze, 2002). Vision is considered better at discriminating details of spatial geometry (shape, color etc.) whereas haptics is particularly effective for the detection of texture (Verry, 1998).

Vision involves both the acquisition and processing of visual information by the visual system. Humans see when the lens of the eye focuses objects in the environment onto the retina; light-sensitive membrane at the back of the eye. The retina contains two types of photoreceptive cells: rods and cones (Howard, 1996). Rods are responsible for vision in low light, and cones handle color vision and detail. When light contacts rods and cones, a series of complex chemical reactions occur. This chemical reaction lead to the formation of activated rhodopsin, which causes electrical impulses in the optic nerve (Bruce et al., 2003). The electrical impulses are then transmitted to the brain where it is encoded and interpreted as light. The human visual system is sensitive to only a small fraction of the electromagnetic spectrum. The visible spectrum for humans ranges from about 400 nm to 700 nm as shown in Figure 2.1. The wavelengths in the visible spectrum have no intrinsic color; however, humans perceive color as a result of interpretation by our visual system. The human eye can discern differences between 8 and 12 million colors; however, we can reliably recall and identify only 6 to12 colors (Howard, 1996).

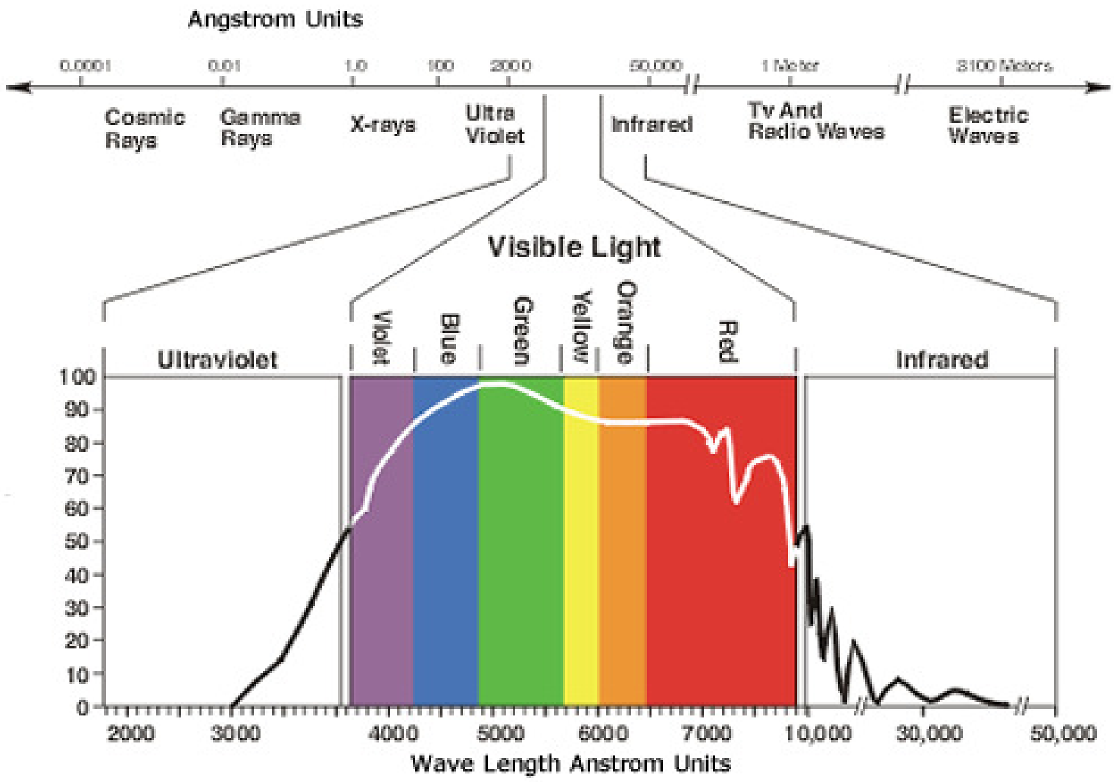

Figure 2.1: Electromagnetic spectrum of visible light for humans [Courtesy of Howard, 1996]

In most human computer interaction, the human receives visual information from the computer through the windows, icons, menus, buttons etc. in graphical user interfaces (GUI); however, there are limited input devices or 'perceptual organs' by which the computer senses the intent of their human users (Quek, 1995). Vision-based interfaces use computer vision (vision as a communication channel) in order to sense and perceive the user and their actions within the context of human computer interaction. The application of computer vision to sense human communication without obstruction has the ability to ensure that the interaction between users and computers is truly natural.

2.2 Auditory Modality

Auditory is the physiological process by which humans perceive sound. It involves the transformation of sound vibrations into nerve impulses in the inner ear which are then transferred to the brain where it is ultimately interpreted as sounds. The human ear can detect frequencies of 20Hz-20 kHz; however, it is most sensitive to those between 1 kHz and 3 kHz. Little or no speech information of value can be extracted above 8 kHz frequency. Similarly, perception of frequencies below 100Hz is tactile in nature and therefore difficult to assess (Truax, 2001).

Usually, sound stimulation enters the ear canal as a sound pressure wave and is converted to vibrations of the middle ear. This vibrations lead to a corresponding motion in the cochlear fluid as a result of the movement of the stapes footplate (Stenfelt et al., 2004). Through this mechanism, sound energy is transduced to electrical nerve energy which is then passed up the auditory nerve to the brain for interpretation. The transduction is accomplished by the displacement of tiny hair cells along the basilar membrane as the membrane moves differently to sounds of different frequencies (Wickens et al., 2004).

Though auditory stimuli last for a short period of time, it is very effective in attracting user's attention and providing information on changing circumstances to users as is the case in alarms. Auditory modality, thus, allows the artificial modification of sound characteristics such as pitch and tone, to convey information to users without the need to focus the sound at the location where it is presented (Leinonen et al., 1979). Another reason that accounts for the effectiveness of auditory modality in attracting users' attention is due to the fact that it is able to provide information beyond the reach of either visual or haptic modality. Sound has the unique ability to provide information from all the directions to users. As a result, it is possible for people to listen and hear actions even when they cannot see the source.

In spite of the inherent expression capabilities and advantages that auditory modality offers, its application in human-computer interfaces has been limited. For example, overlapping auditory information with visual information in a human-computer interface will reinforce users' ability to recollect and, therefore, prolong the time associated with memory fading due to the limitations of the working memory (Leinonen et al., 1979). Two types of audio messages that are used to present information to users in human-computer interface are auditory icons and earcons.

Auditory icons are sounds designed to convey information about events by analogy to everyday sound-producing events (Bjur, 1998; Gaver, 1989). In other words, auditory icons are sounds from the everyday environment used in human-computer/machine interfaces to help users understand what kind of information they are dealing with. For example, the sound of objects crashing into a trashcan may be used to denote the deletion of a file; as a result, auditory icons have the advantage of being understood by users without learning or memorization. Earcons on the other hand, are abstract tones that are used in structured combinations to create sound messages to represent parts of an interface (Brewster, 1997). For example, a tone may be used in an interface to represent an invalid operation by the user. However, since earcons have no intuitive relationship with sounds and objects they represent, users have to learn and recall their associations. Other benefits of auditory channel when combined with other senses in a human machine interface as summarized by Kramer (1994), are shown in Table 2.2.

Table 2.2: Advantages of auditory channel in human machine system [Courtesy of Kramer, 1994]

Quality	Advantage
Non-intrusive enhancement	Augments visual displays without interfering with existing tools and skills
Increase in perceived quality	Affordable and easily
Superior temporal quality	Time series data. Shorter duration events can be detected with auditory displays
High dimensionality	Adds to and exceeds dimensionality of visual and haptic modalities
Engagement	Decreases learning times, reduces fatigue, and increases enthusiasm
Complementary pattern recognition capabilities	Provides the opportunity to bring new and different capabilities to the detection of relationships in data
Inter-modal co-relation	Reinforcement of sensed experiences, veridical representations
Enhanced realism	Immersive, interactive interfaces become more realistic
Synsethesia	Replacement of inappropriate or insufficient cues from other sensory channels
Enhanced learning and creativity	Provides a representation modality suited to the student's learning style, encourages fresh interpretations techniques, imagined extensions of data when auditing
Lower computational requirements	Efficient use of CPU and memory resources for display tasks

2.3 Olfaction Modality

Olfaction, the sense of smell is a direct sense. Humans can distinguish more than 10,000 different smells or odorants usually detected by specialized olfactory receptor neurons that line the nose (Buck & Axel, 1991). The olfactory receptor neurons are out in the open in the nostrils and, therefore, come into contact with the air. They have hair-like projections called cilia which increase their surface area. When volatile materials give off molecules or odorants, air current sweeps these odorants up through the nostrils until the molecules hit the olfactory epithelium. The airborne molecules stimulate the olfactory receptor cells, bind to the cilia and send electrical impulses to the brain. The brain then interprets patterns in electrical activity as specific odors and the olfactory sensation is perceived as smell.

Smell is man's first response to stimuli, and may alert us to dangers before other senses do. For example, gas leak may be smelled before there is an explosion; fire may be smelled before there are flames, and we may recoil before we taste rotten food. In spite of smell being man's first response to stimuli, it has received little application in human-computer interaction, in part due to the challenges involved with odor generation and control of breathing space.

Garcia-Ruiz et al. (2008) defined an olfactory interface as one that employs one or more natural or artificially-created odors in a computer interface, with a purpose of assisting the human user. This could be achieved through the use of olfactory icons, which are computer-generated scent that conveys meaningful information to users, and must be semantically and environmentally related to the information to be conveyed

(Kaye, 2004) . For example, a computer display showing a virtual environment of wildfire may release the smell of burning forest and will augment the user's immersion experience. Similarly, Shoeib et al. (2006) reported that the United States military in conjunction with the University of Southern California have developed a realistic virtual reality olfactory interface simulator that integrates smell to enhance training in war zone. In this simulator, soldiers wear an electronic collar which generates scents and odors through wireless network according to the activities performed and events generated in the virtual environment. For example, when soldiers shoot guns in the virtual environment, the electronic collar generates the scent of gun powder, which the soldiers can perceive. This enhances the soldiers' perception of real war zone and makes training more effective. Another olfactory interface "Dollars & Scents" developed by Kaye (2001) mimics the changes in the stock market by releasing scents such as roses when the stock market is going up and lemons when the stock market is going down into the air. This provides a scent reminder by allowing users to create smell alarms.

Thus, in principle, olfactory cues can be used to support such functions as sensory substitution and to convey high-level assessment of situations (such as alerting users to the presence life threatening stimuli), mood manipulation, increasing vigilance, decreasing stress, and improving the retention and recall of learned material (Sarter, 2006). However, in doing so, care must be taken to avoid potential allergy and nausea reactions. This is further complicated by the difficulty in creating and delivering smell in human-computer interfaces, and the fact that smell is good for slow changing events but not so good for fast changing events often encountered in human-computer interaction.

2.4 Taste Modality

Taste (gestation) is the sensation that results when taste buds in the tongue and throat convey information about the chemical composition of a soluble stimulus. In other words, taste is the sense that distinguishes the sweet, sour, salty, and bitter qualities of dissolved substances in contact with the taste buds on the tongue. Humans detect taste with taste receptor cells clustered in the taste buds found on the surface of the tongue (Erickson, 1982). Each taste bud has a pore that opens out to the surface of the tongue enabling molecules and ions taken into the mouth to reach the receptor cells inside. Humans like many other vertebrates, combine the sense of taste with the less direct sense of smell in the brain's perception of flavor.

2.5 Haptic Modality

Haptics refers to sensing and manipulation through touch. The origin of the word haptic can be traced back to the Greek words: *haptikos* meaning "able to touch" and *haptesthai* which translates to "able to lay hold of" (Katz & Krueger, 1989; Révész, 1950). However, today it is used broadly to encompass the study of touch and the human interaction with external environment via touch. More commonly, the word "haptic" or "haptics" refers to the capability to sense a natural or synthetic mechanical environment through touch and includes kinesthesia (or proprioception), the ability to perceive one's body position, movement and weight (Hayward et al., 2004). The field of haptics is inherently multidisciplinary and draws from many disciplines, including biomechanics, neuroscience, psychophysics, robot design and control, mathematical modeling and simulation, software engineering, and systems control among others.

The term haptics has been used by psychologists who study how people use their hands to sense and manipulate objects since the early part of twentieth century (Salisbury, 1995). Although humans interact with our surroundings through five sensory channels: sight, sound, taste, smell, and touch, it is only the sense of touch that enables humans to modify and manipulate the world around them (McLaughlin et al., 2002). Most of the information that humans gain by means of touch comes by way of the hand, which is both a perceptual and manipulative organ.

The recent explosion in computer technology and the need for better and intuitive ways for humans to interact with machines and computer-generated virtual environments has led to increased interest in haptics. It promises profound changes to the way humans interact with machines by allowing users to have the sensation of '*feel*' through the provision of force feedback (simulating object hardness, weight, and inertia) and/or tactile feedback (simulating surface contact geometry, smoothness, slippage, and temperature (Jacobson et al., 2002). Haptic interface devices share the unparalleled ability to provide for simultaneous and bi-directional information exchange between a user and a machine/computer as shown in Figure 2.2 (Minogue & Jones, 2006).

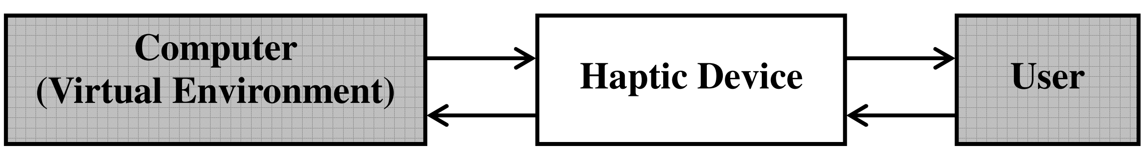

Figure 2.2: Illustration of the simultaneous exchange of information between user and machine unique to haptic interfaces [Courtesy of Minogue & Jones, 2006]

Tactile stimulus is conveyed to the brain in one of two ways: tactile information or kinesthetic information. Tactile information refers to the sense of natural contact with an object while kinesthetic information refers to the sense of position and motion. Srinivasan (1995) defined haptics into three sub-groups: human haptics, machine haptics and computer haptics. Human haptics is the study of human sensing and manipulation through touch. Machine haptics is the complimentary study of the design and construction of machines and includes the development of technology to augment haptic communication between humans and machines as illustrated in Figure 2.3 (Srinivasan, 1995). In this illustration, the human senses and controls the position of the hand, while the machine exerts forces to simulate contact with virtual objects. The human input system includes human senses that the operator uses to receive feedback from the machine and the environment such as the eyes for sensing visual feedback, ears for sensing auditory feedback or skin for sensing vibrations.

Both the human and machine systems have sensors in the form of nerve receptors or encoders, processors in brain or computer, and actuators in muscles or motors. Computer haptics involves the development of algorithms and software needed to generate and render touch and feel into objects in virtual environments. A detailed description of haptics, its applications and an empirical study aimed at identifying the appropriate force feedback for use in haptic control excavator are provided in Chapter 4 of this dissertation. Further, a conceptual model of the human-excavator interaction is described in Section 5.3.

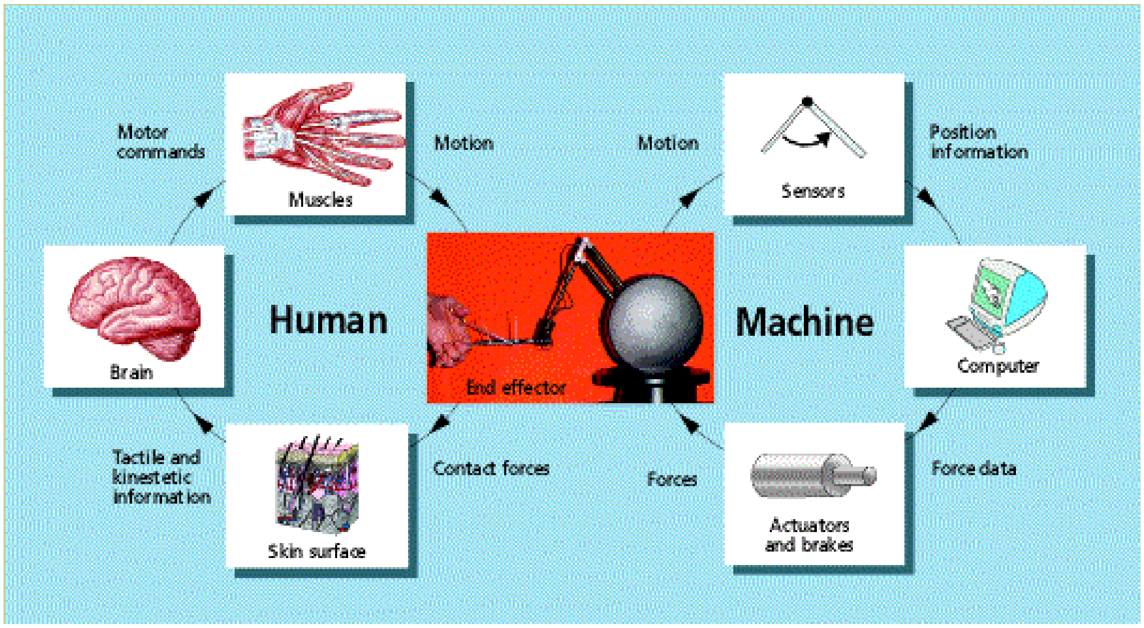

Figure 2.3: Machine haptics interface with human [Courtesy of Srinivasan, 1995]

By using special input/output devices: joysticks, stylus, data gloves, or other devices, users can receive feedback from computer applications in the form of 'feel' sensations in the hand or other parts of the body (Jacobson et al., 2002). In combination with a visual display, haptics technology can be used to train people for tasks requiring hand-eye coordination to reduce errors and improve performance.

2.6 Haptic User Interface

A haptic user interface is an interface that uses computer-controlled mechanism to allow users to interact with systems/machines through the sense of touch. Haptic provides an intuitive interface between man and machine, and requires little training and a working style most like that used by humans to interact with their environment and objects in day-to-day life. In other words, the human interacts with elements of his/her task by looking, holding, manipulating, listening, and moving, thus, using as many of his/her natural skills as appropriate, or can reasonably be applied to a task (Stone, 2001). Haptic user interfaces vary significantly in their complexity, and may range from devices with passive haptic displays or simple vibrotactile haptic feedback to complex systems with dynamic haptic feedback and sensing of finger, hand, head or body movement (Bjelland & Tangeland, 2007). Haptic user interfaces have wide range of applications that range from surgical devices in medicine to aviation, gaming and virtual reality industries, though their use in commercial products is low due to the technical challenges of their implementation.

2.7 Theories of Multimodal Human Machine Interfaces

Most of the advantages of multimodal human machine interface designs are rooted in the theory of cognitive psychology and human computer interaction studies (Dumas et al., 2009) . Specifically cognitive load theory, gestalt theory, Baddeley's model of working memory as well as Wicken's multiple resource theory are among the theories most often used to elucidate multimodal human machine interfaces (A. Baddeley, 1992; Wickens, 2002). As outlined by Dumas et al. (2009), research in cognitive psychology have shown that,

(1) Humans are able to process multiple modalities partially independently, and therefore, presenting information with multiple modalities increases the humans working memory.

(2) Humans have the tendency to mimic interpersonal interaction habits during multimodal interaction with machines/systems.

(3) Human performance is improved when interacting multimodally due to the way that human perception, communication, and memory systems function.

Thus, some theories of cognitive psychology which help in explaining some of the benefits of multimodal human machine interface design are discussed below.

2.7.1 Cognitive Load Theory.

Cognitive load theory proposes that working memory is limited and, therefore, if learners are bombarded by information and, if the complexity of their instructional materials is not properly managed, then it will result in a cognitive overload. This cognitive overload impairs schema acquisition, later resulting in a lower performance (Sweller et al., 1990) . Once learners have acquired a schema, those patterns of behavior (schemas) may be practiced to promote skill automation (Anderson, 1982) but expertise occurs much later in the process, and is when a learner automates complex cognitive skills usually via problem solving (Shiffrin & Schneider, 1977). Schema acquisition is the ultimate goal of cognitive load theory. Anderson's framework proposes initial schema acquisition occurs by the development of schema-based production rules, either by developing these rules during practice or by studying examples.

Sweller et al. (1990) also argued that since the working memory is the primary limitation in learning, presenting information in multiple modalities rather than single modality could help expand the processing capabilities by increasing the effective working memory. In other words, cognitive load theory assumes a limited working memory in which all conscious learning and thinking occurs, and an effectively unlimited long-term memory that holds a large number of automated schemas that can be brought into working memory for processing (Dumas et al., 2009). Oviatt (2006) applied the cognitive load theory by testing different educational interface designs and found that user interface designs that minimized cognitive load freed up mental resources and improved students' performance.

2.7.2 Gestalt Theory.

Gestalt theory refers to the form-forming capability of the human senses, particularly with respect to the visual recognition of figures and whole forms instead of just a collection of simple lines and curves. The phrase "The whole is greater than the sum of the parts" is often used to explain Gestalt theory (Koffka, 1999; Koka, 1935; Reiser, 1936). It was applied to visual perception by Wertheimer, Kohler and Koffka who founded the so-called gestalt approaches to form perception, with the goal to investigate the global and holistic processes involved in perceiving structures in the environment (Sternberg et al., 2009) . More specifically, the Gestalt theory explains how humans perceive groups of objects and how these perceived parts of objects form whole objects.

The Gestalt theory has many principles/concepts with wide applications in human computer interactions. A brief summary of key concepts is provided below.

Similarity - occurs when objects look similar to one another and, therefore, is perceived as a group or pattern. For example, in design similarity could be broken in order to attract user's attention by using highlighting, underlining, sound, flashing or animation.

Continuation - occurs when the eye is compelled to move from one object to another. It is the eye's instinctive action to follow a direction derived from the visual field (Fultz, 1999; Koffka, 1999).

Closure - occurs when an object is incomplete or a space is not completely enclosed. However, if enough of the shape is shown, then, users will be able to perceive the whole by filling in the missing parts. For example, a reader may be able to read the word '*stud-nt*' as student though one word is missing. This is due to the fact that most times users do not read the individual letters but rather the complete words.

Proximity - occurs when elements are placed close together that they are perceived as a group. Users are able to mentally organize closer elements into a coherent object, because closely spaced elements are assumed to be related and those further apart are assumed to be unrelated (Fultz, 1999).

Ground/Figure – occurs when the eyes differentiate an object from its surrounding area. A form or shape is naturally perceived as foreground while the surrounding area is perceived as background. For example, two different foreground colors may result in a user perceiving two different things from the same image and, therefore, a good balance between figure/ground relationships can make the perceived image clearer and add interest and detail to the image (Chang et al., 2002).

Oviatt (2003) demonstrated that a number of human behaviors could be successfully predicted by design of map-based pen/voice interfaces using Gestalt principles. They observed that users consistently followed a defined multimodal integration pattern (sequential or simultaneous) during error handling and became entrenched in their patterns rather than change their behavior. Further, they observed that Gestalt theory was accurate in predicting that a dominant number of subjects applied simultaneous integration over sequential integration.

2.7.3 Baddeley's Model of Working Memory.

Baddeley and Hitch proposed their tripartite working memory model as an alternative to the short-term store in 1968. This model was later expanded upon by Baddeley and has become the dominant view in the field of working memory. The original short-term model by Baddeley & Hitch (1974) was composed of three main components: the central executive which acts as supervisory system and controls the flow of information from and to its slave systems, the phonological loop, and the visuo-spatial sketchpad. The slave systems are short-term storage systems dedicated to content domain (verbal and visuo-spatial, respectively). Baddeley (2000) added a third slave system, the "episodic buffer" to his model. The episodic buffer is dedicated to linking information across domains to form integrated units of visual, spatial, and verbal information with time sequencing (or chronological ordering), such as the memory of a story or a movie scene. The episodic buffer is also assumed to have links to long-term memory and semantic meaning (Baddeley, 2000). Though the slave processors are coordinated by a central executive, they function largely independently in terms of lower level modality processing (Dumas et al., 2009). Baddeley's model of working memory is shown in Figure 2.4. Baddeley & Hitch (1974)'s argument for the distinction of two domain-specific slave systems in the older model was derived from experimental findings with dual-task paradigms.

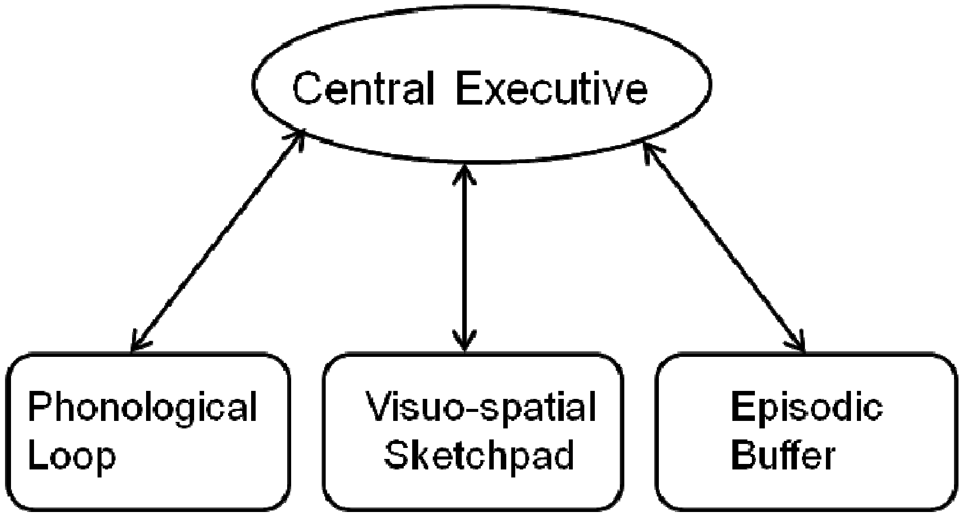

Figure 2.4: Baddeley's model of working memory

Performance of two simultaneous tasks requiring the use of two separate perceptual domains (i.e. a visual and a verbal task) was nearly as efficient as performance of the tasks individually. In contrast, when a person tried to carry out two tasks simultaneously that used the same perceptual domain, performance was less efficient than when performing the tasks individually. Therefore, performance is improved when humans interacted with two modalities that can be processed simultaneously in separate resource pools. The Baddeley model's strength is its ability to integrate large amount of findings from work on short-term and working memory.

2.7.4 Resource Theory Model.

The resource theory models developed by early researchers (Kahneman, 1973; Navon & Gopher, 1979; Norman & Bobrow, 1975; Wickens, 1980; Wickens & Liu, 1988) conceptualized attentional resources as commodities or pools of energy to be spent on task performance. First, single resource theory proposed by Kahneman (1973) argued that there was a single pool of limited capacity available for a variety of tasks, and that performance depends upon the degree to which that capacity is allocated to tasks. However, Navon & Gopher (1979) proposed a multiple resource theory that argued that the human is a multiple processor, and that each processor may have its own capacities and that each capacity may be shared by several processors. Wickens (1984) extended the idea of multiple resource theory and argued that, there exist multiple attentional resources that are sometimes separate from one another which can be tapped simultaneously. He proposed separate pools of resources for information processing codes (spatial and verbal), different input modalities (visual, auditory, etc), different stages of information processing (encoding, central processing and responding), different response type (motor or verbal), and argued that these resources could be utilized separately or jointly depending on the information processing needs of the tasks at hand. These resource models are further described in Section 3.2.

CHAPTER 3

INTERFERENCE IN MULTIMODAL HUMAN-EXCAVATOR INTERFACE

3.1 Background

A multimodal human-machine interface refers to human-machine interaction in which the user interacts with the application by using two or more input/output modalities. Thus, a human/machine system or interface is multimodal if it supports multiple human sensing modalities such as vision, auditory, touch or smell in the interaction. Multimodal interfaces represent a shift and a new paradigm in human-machine interface design from the traditional graphical interfaces to interfaces that make use of the natural human-to-human interaction characteristics. The advent of multimodal interfaces based on recognition of human speech, gaze, gesture, touch and other natural behavior represents only the beginning of progression towards computational interfaces capable of human-like sensory perception (Thiran et al., 2009).

A well-designed multimodal system integrates complementary modalities to yield a highly synergistic blend in which the strengths of each mode are capitalized upon and used to overcome weaknesses in the other (Oviatt, 1999). The user, therefore, has multiple modalities available for input/output interaction with the system. The purpose of using multiple sensing modalities in human machine interfaces is to design systems/interfaces that are easy to use, efficient, flexible, transparent, and provide a highly expressive mode of interaction between humans and computers/machines. Further, humans naturally interact with their environment multimodally. For example, humans can speak about an object, touch the object, look at the object, point at the object, and smell the object concurrently.

In everyday activity, humans interact with our environment through coordinated visual-haptic perceptions. For example, humans use both visual and haptic senses jointly in order to reach toward and grasp objects. Through the joint use of visual and haptic senses, that is looking at an object while handling it, humans are able to evaluate information about the shape, weight, texture, etc., that otherwise would not have been possible using either visual or haptic sense independently. The human brain as well as the brain of other dexterous primates contains specialized bimodal visuo-tactile neurons that coordinate information from vision and touch (Graziano & Gross, 1993; Leinonen et al., 1979). According to Keehner (2008), the human memory, at a higher level of representation, uses similar parameters to code information from the eyes and from the hands, which allows the two sensory modalities to share common spatial reference frames. This ensures that memory representations arising from the two modalities are similar for objects regardless of whether they are perceived through vision or through touch. As a result, spatial cues presented in one modality can speed reactions to spatial cues presented in the other modality (Keehner, 2008).

Multimodal interaction uses different modalities, visual, audio, and tactile feedback, to engage human perceptual, cognitive, and communication skills and help users to understand better what is being presented, which ultimately lead to more engaged user interaction and improve performance. Using multiple modalities in human-computer/machine interaction will make the environment resemble that in which humans naturally operate and result in a more efficient and effective interaction. Sharma et al. (1998) identified four issues that need to be considered in multimodal human machine interfaces as

(a) Why integrate multiple modalities?

(b) Which modalities to integrate?

(c) When to integrate multiple modalities?

(d) How to integrate multiple modalities?

3.1.1 Rationale for Multimodal Human-Machine Interface

Since humans interact naturally with their environment multimodally (i.e. see, touch, hear, smell, taste), it is envisaged that multimodal human-machine interfaces will migrate the natural habits used by humans to communicate with one another into the human-computer interaction (see Figure 3.1). Further, most human machine/computer interaction devices are practically unnatural and cumbersome, and rely on devices such as mouse, joystick, keyboard, pedals, etc., which effectively limit the ease with which the user can interact with the machine/computer (Sharma et al., 1998). Also, multimodal human-machine interface offers the user freedom to use a combination of modalities, or to switch to a better-suited modality, depending on the specifics of the task or environment at hand (Oviatt et al., 2000). For example, individual input modalities may be well suited in some situations, and less ideal or even inappropriate in others, and therefore, the choice of appropriate modality should be task dependent and requires careful consideration in the design of multimodal system.

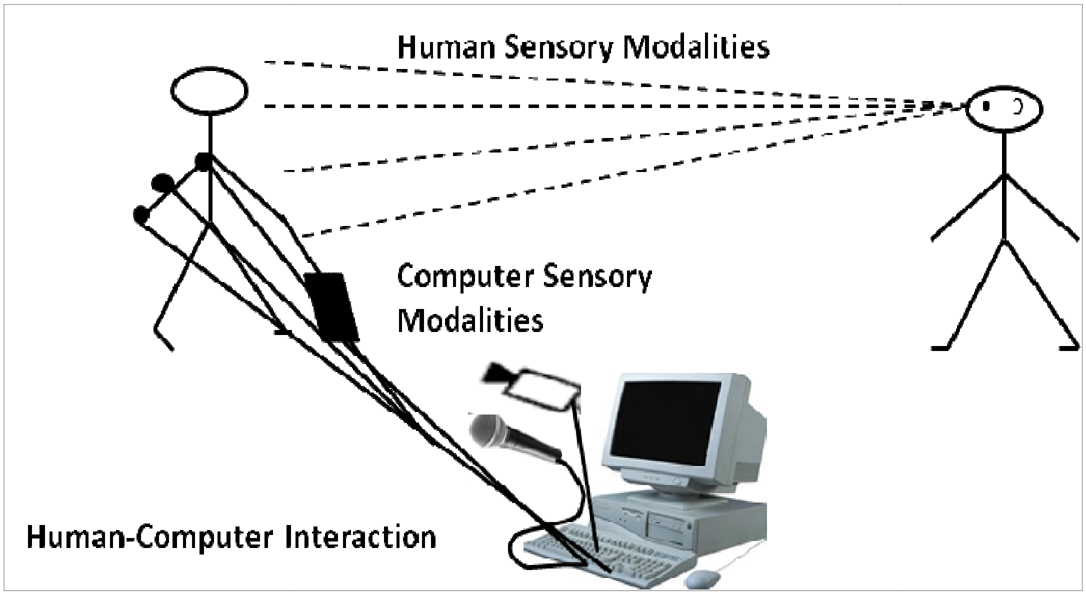

Figure 3.1: Human-to-human interaction and human-to-computer interaction
[Courtesy of Sharma et al., 1998]

Another rationale for integrating multiple sensory modalities into multimodal human machine interfaces is that the superior colliculus of the brain which receives and transmits signals from the cerebral cortex is multisensory. Therefore, there is a strong indication that using multiple modalities in human-machine interfaces would be desirable, if the goal is to incorporate the *naturalness* of human communication into human computer interaction (Sharma et al., 1998). Furthermore, it is statistically advantageous to combine multiple observations from the same source because improved estimates are obtained using the redundant observations and that multiple types of sensors may increase the accuracy with which the quantity can be observed (Hall & Llinas, 1997). A summary of potential benefits offered by multimodal human-machine interfaces as outlined by Oviatt et al. (2000) is provided below.

(1) They permit the flexible use of input modes, including alternation and integrated use.

(2) They support improved efficiency, especially when manipulating graphical information.

(3) They lead to enhancement in system robustness.

(4) They can support greater precision of spatial information than a speech only interface, since pen input can be quite precise.

(5) They give users alternatives in their interaction techniques.

(6) They lead to enhanced error avoidance and ease of error resolution.

(7) They accommodate a wider range of users, tasks, and environmental situations.

(8) They are adaptable during continuously changing environmental conditions.

(9) They accommodate individual differences, such as permanent or temporary handicaps.

(10) They can help prevent overuse of any individual mode during extended computer usage.

In addition to the benefits outlined above, recent research indicates that humans may process information faster and better when it is presented in multiple modalities (Turk & Kölsch, 2004). A well-designed multimodal interface that allows user flexibility can potentially leverage people's natural ability to use modes accurately and efficiently (Oviatt et al., 2000). Multimodal interfaces are, therefore, desirable in human-machine interaction because they support synergy, redundancy, disambiguation, and offer increased bandwidth for information exchange between humans and machines. However, making a system multimodal by just adding a further modality to the system may not necessarily lead to improvement in the system as this may increase operator's cognitive load due to the increased degrees of freedom. In some cases the different modalities may actually interfere with each other and may have a negative impact on system performance.

3.1.2 Modalities Selection

Humans perceive their environment through visual, touch, hearing, smell and taste senses and interact with it by using their actuators such as hands, body, face, voice, etc. (Sharma et al., 1998). Modalities for human sensing and actions are shown in Figure 3.2. In human-to-human interaction, actuator actions of one human are perceived by senses of the other human in the environment, whereas in human-computer interaction, the computer perceives actions of humans. Thus, in order for human-computer interaction to be as natural and effective as possible, it is important that the computer is able to interpret all human actions such as hand, gaze, body, speech, gestures, etc. The richness of the interaction in human-to-human interaction can be attributed to the ability of humans to sense changes in facial expressions, gestures, body movements and other human expressions that are unavailable in human-to-computer interaction. Some computer sensory modalities are analogous to human sensory modalities; however, computers possess other sensory modalities that humans' lack, such as the ability to accurately estimate the position of the human hand through magnetic sensors and measure subtle changes in the electric activity of the human brain (Sharma et al., 1998).

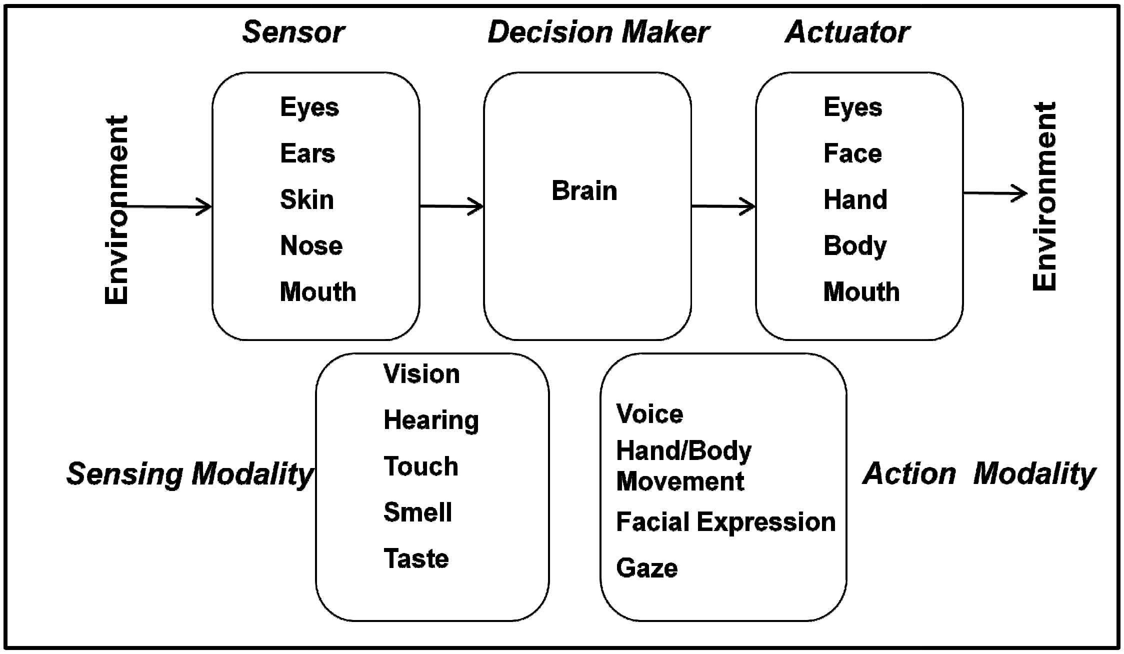

Figure 3.2: Modalities of human sensing and action [Courtesy of Sharma et al., 1998]

Figure 3.3 shows human action modalities and computer sensing modalities and how they are related to each other. For example, a human action modality such as speaking may be interpreted by more than one computer sensing modality such as video and audio. Computer-sensing modalities such as position and motion sensing, audio sensing, visual sensing, tactile and force feedback sensing, as well as neural sensing could be explored, and if sufficiently developed could be integrated with human sensing modalities to create efficient and effective human machine interfaces. Multiple human actions such as facial expressions and hand or eye movement can be sensed through the same devices and used to infer different information.

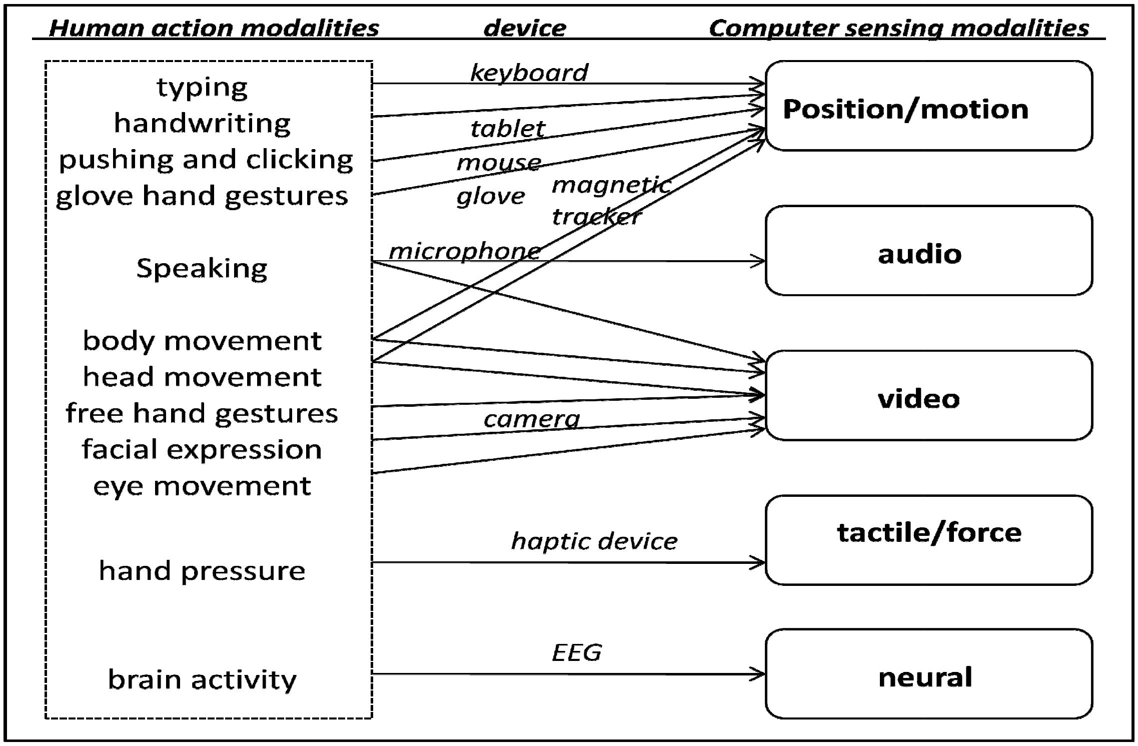

Figure 3.3: Mapping of different human-action modalities to computer-sensing modalities for human computer interaction [Courtesy of Sharma et al., 1998]

3.2 Conflicts in Multimodal Interfaces

As stated by Oviatt et al. (2000), making a system multimodal by just adding a further modality to the system may not necessarily lead to improvement in the system and may instead increase operator's cognitive load due to the increased number of processing channels. In some cases the different modalities may actually interfere with each other and may have a negative impact on system performance.

Three of the most influential theories that are used to explain the decrease in operator performance as task modalities increase are capacity sharing, bottlenecks (task switching), and cross talk (Pashler, 1994). Capacity sharing models are the most widely accepted theory for dual-task interference and assumes that humans have limited mental resources that must be shared among tasks during processing, so that whenever more than one task is performed, there is less capacity for each individual task and performance is impaired (Pashler, 1994). Bottleneck (task switching) model of task interference assumes that parallel processing may not be possible for certain mental operations and, therefore, such operations may require single mechanism to be dedicated for a period of time. When two tasks that require the same mechanism are to be performed at the same time a bottleneck results and one or both tasks will be impaired. The third interference models are the cross talk models in which interference do not depend on the sort of operation being carried out but on the content of information that is actually being processed, such as the type of sensory inputs present, and type of responses being produced.

For capacity sharing models, the decrease in performance when multiple modalities (resources) are available could be traced to the theories of single and multiple resources. The resource concept is founded on the underlying assumption that the human operator has limited capacity for processing resources that may be allocated to task performance (Wickens & Liu, 1988). Therefore, performing two tasks simultaneously require timesharing the scarce resource which may lead to one or both having fewer resources that required and may deteriorate performance.

As described by Szalma & Hancock (2002) , most of the early researchers in resource theories (Kahneman, 1973; Navon & Gopher, 1979; Norman & Bobrow, 1975; Wickens, 1980) conceptualized resources as commodities or pools of energy to be spent on task performance. Resources were described with either economic (supply and demand) or thermodynamic (tank of liquid to be divided among tasks) metaphors. In the economic model, performance on one or more tasks suffered when the resource demands of the tasks exceeded available supply, while in the thermodynamic, resource was viewed as a tank of liquid to be divided among several tasks, and that under stressful conditions the amount of resources available is depleted and performance suffered. The problem with these over-simplified non-biological models was that it failed to include the complex, dynamic and adaptive characteristics of humans.

One of the first authors to postulate the human processing capacity as a resource was Kahneman (1973), who proposed that there is a single pool of limited capacity available for a variety of human tasks, and that performance on tasks depends upon the degree to which a resource capacity is allocated to the particular task. Thus, the single resource theory argues that operators have access to mental resources which they can strategically allocate to multiple tasks, and that task interference is dependent on task difficulty. The amount of capacity available is assumed to be limited and a function of arousal level. An allocation strategy determines how much processing capacity each task receives, and the strategy adopted is influenced by characteristics of the individual and motivational factors (Szalma & Hancock, 2002).

3.2.1 Multiple Resource Model.

The multiple resource theory, first proposed by Wickens (1980) argues that the human is a multiple processor, that each processor may have its own capacities and that each capacity may be shared by several processors. Wickens (1981) extended the idea of multiple resources and proposed the existence of multiple attentional resources that are sometimes separate from one another and which can be tapped independently or jointly.

He proposed separate pools of resources for information processing codes (spatial vs. verbal), different input modalities (visual vs. auditory vs. haptic), different stages of information processing (encoding/central processing vs. responding), and different response type (motor or verbal), and argued that these resources could be utilized separately or jointly depending on the information processing needs of the tasks at hand.

Wickens' three-dimensional representation of the structure of multiple resources is shown below. Wickens' model identified four important categorical and separate dimensions that account for the variance in time-sharing performance. These are the stages of processing, the codes of processing, modalities of input, and visual channels as shown in Figure 3.4. In the processing stages, perceptual and cogintive tasks (involving working memory) use resources that are different from action/response tasks. The code processing stage indicates that spatial activity and verbal activity use different attentional resources. Research has also shown that in the modalities dimension, processing of visual, auditory and haptic stimuli rely on separate attentional resources at the perceptual stage; however, at the central processing stage, common attentional resource may be used to process them (Smith & Buchholz, 1991). Thus, using Wicken's multiple resource theory, the level of disruption or inteference between two tasks that are time-shared can be predicted, and can be used as a guide by designers to make decisions on whether to use voice or manual control, use auditory or visual displays, etc. Further, Wickens (2002) proposed the existence of separate resources which are both limited in capacity and allocatable amongst different tasks in human information processing system.

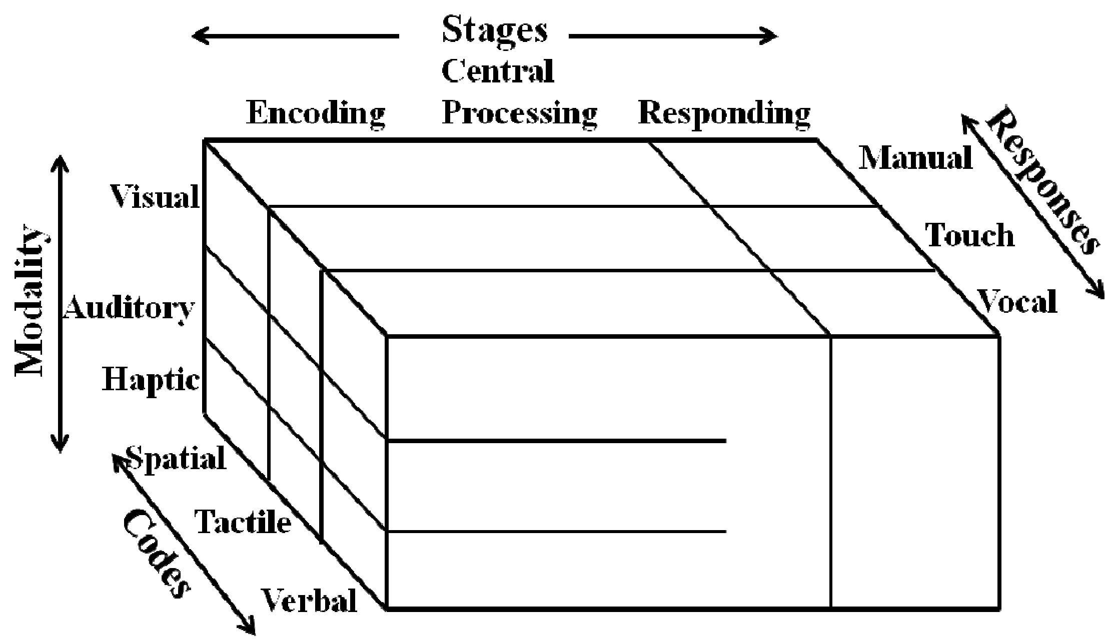

Figure 3.4: Three dimensional resource model [Courtesy of Wickens, 1981]

Depending on the nature of the tasks, these attentional resources may process information sequentially if the different tasks require the same pool of resources, or can be processed in parallel if the task requires different resources.

Thus, tasks that require separate attentional resources will be time shared efficiently without significant cross-task interference. However, performance on tasks that require common attentional resources will depend on how attentional resources are allocated to the tasks due to interference. In other words, humans have limited cognitive resources, therefore, when an operator performs two or more tasks that require attentional resources from a single resource pool, task performance will be impaired since demand for attentional resources exceeds supply. The importance of the concept of multiple resource theory lies in its ability to predict dual task interference levels between concurrently performed tasks, to be consistent with the neurophysiologic mechanisms underlying task performance, and to account for variability in task interference (Wickens, 2002). Thus, multiple resource theory predicts greater task interference when multiple tasks compete for limited and overlapping resources and when task difficulty is increased, as opposed to tasks that are easy, or tasks that draw on non-overlapping resource pools (Horrey & Wickens, 2003; Navon & Gopher, 1979).

Several empirical studies have shown that the utilization of the separate resources outlined in multiple resource theory lead to improved user performance on information processing and recall tasks compared to single modality presentation, and that users find it easier to attend to information displayed using multiple modalities than unimodal systems (Parkes & Coleman, 1990; Wickens, Sandry, & Vidulich, 1983). For example, Wickens (1976) showed that, individuals performed better when task responses are distributed across manual and auditory inputs compared to when two manual or auditory responses were required. Similarly, Wickens & Liu (1988) showed that individuals performed better in a manual tracking task while simultaneously responding verbally to a secondary tone identification task, than when the secondary task required a manual response.

Although the human sensory systems are traditionally thought of as distinct modes of resources, cross-modal interactions are increasingly being recognized for playing a vital role in human perception (Sathian & Zangaladze, 2002). A great ability of the human vision system is the capability for rapid and seemingly effortless recognition of objects despite variations (such as changes in viewpoint or illumination) in the sensory information about the object. In visual information processing, objects of interest attract the eye and initiate a saccade that moves the object of interest from peripheral vision to the fovea for detailed assessment (Woods & Newell, 2004). Here, depth cues are used to deduce object's size whereas other information such as texture and coloring as used to provide material composition (Woods & Newell, 2004). The next section describes an eye tracking study that was conducted to assess conflict and interference between visual, audio and haptic modalities in an excavator interface, to what extent these conflicts impact performance, and how the impact of such interference could be minimized in the human-excavator interface.

3.3 Using Eye Tracking to Investigate Multimodal Interference

Although eye tracking is a relatively new technique in Human-Computer Interaction (HCI), it has been established as a viable tool for usability assessment (Benel, Ottens, & Horst, 1991). Eye tracking is a technique whereby an individual's eye movements are measured so that the researcher knows both where a person is looking at any given time and the sequence in which their eyes are shifting from one location to another (Poole & Ball, 2005). It involves the monitoring and application of eye movements to user interfaces: both for analyzing interfaces, measuring usability, and gaining insight into human performance, as well as an actual control medium within a human-computer dialogue (Jacob & Karn, 2003).

The structure of the human visual system enables high resolution vision to occur in only a small region, as a result, humans have to adjust and focus their gaze to the location from which visual information needs to be collected at any given moment. Therefore, an individual's eye movements' data provide information about the nature, sequence and timing of cognitive operations. In other words, what a person is looking at is assumed to indicate the thought "on top of the stack" of cognitive processes (Just & Carpenter, 1976), thus this eye-mind hypothesis means that eye-movement recordings can provide a dynamic trace of where a person's attention is being directed in relation to a visual display. By providing information about the nature and timing of mental processes, eye movement information could be useful in systems that attempts to facilitate human computer interaction by giving the computer more information about the user's cognitive activities (Rudmann et al., 2003).

The concept of using eye movements to predict users' thought processes pre-dates the widespread use of computers by almost 100 years, and many different methods have been used to track eye movements since the use of eye tracking technology was first pioneered in reading research (Rayner & Pollatsek, 1989). Some initial eye tracking methods were quite invasive and involved direct mechanical contact with the cornea, however, this has evolved over the years to current non-invasive eye tracking techniques, which use light reflected from the cornea (Jacob & Karn, 2003). Further, eye movement research and eye tracking has flourished with advances in both eye tracking technology and psychological theory to link eye tracking data to cognitive processes. A lot of research work carried out in psychology and physiology have focused on exploring how the human eye operates and what it can reveal about perceptual and cognitive processes.

The most common commercial eye-tracking systems in use today for human computer interface research that use human subjects are video-based pupil/corneal reflection eye-tracking systems. These eye-trackers rely on video localization of the pupil in conjunction with infrared illumination. The infrared illumination reflects off the cornea and the location of the corneal reflections are detected and used as a benchmark to gauge the relative position of the pupil. As a result, these eye-trackers are very resilient to subject motion. These commercial eye-tracking systems measure point-of-regard by the "corneal-reflection/pupil-centre" method (Goldberg & Wichansky, 2003), which usually consist of a standard desktop computer with an infrared camera integrated into display monitor, and image processing software to locate and identify the features of the eye used for tracking. When the eye tracker is turned on, infrared light from an LED embedded in the infrared camera is first directed into the eye to create strong reflections in target eye features so they are easy to track. The use of the infrared light ensures that the user is not dazzled with visible light. When the light enters the retina, a large proportion of it is reflected back, making the pupil appear as a bright, well defined disc known as the "bright pupil" effect (Poole & Ball, 2005). Image processing software is used to identify the center of the pupil and the location of the corneal reflection, the vector between them is measured, with additional trigonometric calculations, point-of-regard is measured. It should be noted, however, that though it is possible to determine approximate point-of-regard by the corneal reflection alone, by tracking both the center of pupil and the location of corneal reflection, eye movements can, critically, be disassociated from head movements (Jacob & Karn, 2003).

3.4 Eye Tracking Study to Investigate Conflicts in Human-Excavator Interface

In order to investigate whether there are conflicts or interferences in the human-excavator interface, an empirical eye tracking study was conducted using the excavator simulator. In this study, the Tobii® Eye Tracker was used to measure, record, play back, and analyze users' eye movement on computer screen as they interacted and manipulated the haptic-controlled excavator. Increasing number of researchers are using the Tobii® Eye Tracker because, assessing the allocation of visual attention with conventional methods such as click analysis, questionnaires or simply asking subjects where they have paid attention to, are limited to those processes which are part of conscious reflection and conscious control. Therefore, relying exclusively on methods like those mentioned above may impact validity of results, since attentional processes do not solely depend on conscious human control. They are often controlled beyond subjects' awareness, are therefore not reportable or are simply too fast to be analyzed by mouse movements (Schiessl et al., 2003). More importantly, what an individual looks at is usually a good reflection of the cognitive processes going on in the mind of that individual.

In conducting this research, a task analysis of the sequence of steps required by an operator to successfully complete an excavation task was conducted to help identify the most critical/crucial tasks in excavator operation. Based on the results from the task analysis, the most critical tasks were identified and used to design a set of tasks to be carried by participants while interacting with an excavator simulation. Also, based on the results from the task analysis, the experimental tasks were designed in such as way that auditory, visual and haptic information are simultaneously presented and used by operators to accomplish the required tasks.

Another design issue that was investigated in this study is hand-eye coordination of the excavator operator. Given that most motor control movement of the operator is either initiated or guided by perception, it is critical that the relationship between excavator operator's gaze and hand movement is fully understood. Several studies have been carried out to find how eye movement relates to hand movement. For example, Binsted et al. (2001) studied the temporal and spatial coupling of gaze and hand movement in direct hand pointing task and found similar patterns of hand-eye movement relationships including: (i) that gaze tended to initiated 70ms earlier than hand movement, (ii) that gaze typically makes two saccades to land on target and that the first saccade tended to undershoot, and (iii) that eye gaze stabilizes on target at 50% of total hand response time. The difference between these previous studies and the current study is that, the current study probed hand-eye coordination of excavator operators using haptic-controlled excavator simulator with haptic feedback. This helped in understanding the relation between hand movement and eye movement when operating the haptic-controlled excavator so that potential operator difficulties could be mitigated.

3.5 Methodology

3.5.1 Research Questions.

This empirical study seeks to answer the following research questions.

1) Does conflict exists between sensory modalities (auditory, visual, and haptic) in the haptic-controlled excavator interface?
2) Are these conflicts significant enough to have an impact on the performance of operator-excavator interaction?
3) Do operators struggle to coordinate their hand-eye movement? If they do, how would this affect the efficient operation of the haptic-controlled excavator?

3.5.2 Participants.

Twenty-four students were recruited from the North Carolina Agricultural & Technical State University (NCA&T) to take part in this empirical study. The participants were grouped into two groups: *novices* and *experts*. The novices group consisted of 20 volunteers with no prior knowledge of the haptic controlled excavator simulator, while the experts group consisted of 4 volunteers made up of members of Center for Compact and Efficient Fluid Power Systems (CCEFP) research team at NCA&T who have had the experience of interacting and manipulating the haptic-controlled excavator.

The sample size for this study was determined using the t-test approach since the study was a comparison study. A 0.80 power was assumed, and using the formulation $p = (ES)*\alpha*\sqrt{n/\sigma}$, where p is the power, ES is the effect size, α is the significance level, n is the sample size, and σ is the standard deviation. Using a pilot study, the effect size (ES) was estimated to be 2, and σ was estimated at 25.29 seconds (0.4215 minutes) from a previous study by the author (Osafo-Yeboah et al., 2010). Using a significance level of 0.05, sample size was calculated to be 27. Thus, 27 participants would have given the study a statistical power of 0.8. However, due to resource and time constraints, only 20 participants were recruited for the study with a statistical power of 0.68.

3.5.3 Equipment.

The equipment for this experiment consisted of two Gateway computers, a Tobii® Eye Tracker T60, and a Phantom Omni 5.3 Haptic device. The two Gateway computers ran the excavator simulation program and were connected with the Tobii® Eye Tracker via a local network. Computer number one interfaced with the Phantom Omni and ran the excavator dynamics simulation, while computer number two ran the xPC-target simulation. The excavator simulation graphics were displayed on the Tobii® Eye Tracker connected via a local network to the two Gateway computers. The schematic layout of the equipment setup is shown in Figure 3.5.

The Phantom Omni device sat next to the Tobii® Eye Tracker on the right hand side of participants and had 6 degrees of freedom in total: up-down, left-right, front-back, and a rotating stylus with 3 degrees of freedom. The C++ and MatLab programming that ran the simulation was developed by Mark Elton of Georgia Institute of Technology.

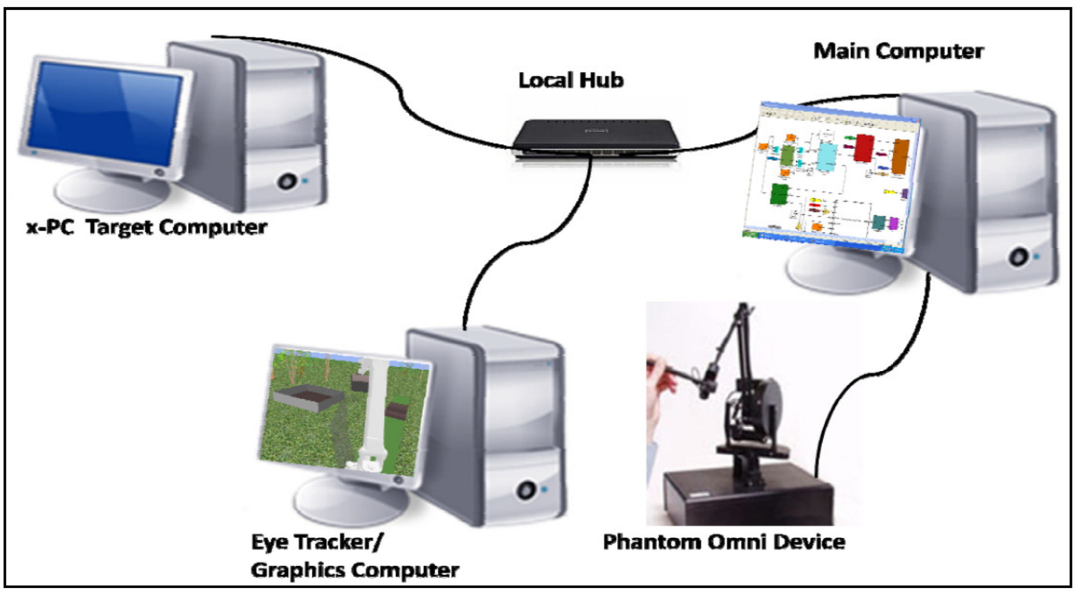

Figure 3.5: Schematic equipment setup of simulation

3.5.4 Experimental Design.

A between-subject design was used in this experiment. The independent variable was expertise with two levels *novices* and *experts*. The dependent variables were task completion time, number of scoops to fill a bin, number of drops outside of the bin. Also, eye tracking data fixation count, fixation length and fixation duration were collected.

3.5.5 Task.

Using the results from the task analysis study described in Chapter 6, participants were asked to perform a series of tasks that required them to move the boom/bucket assembly to the desired location using the stylus of the Phantom Omni device. Next, they had to position the bucket at the work area (trench), then scoop/dig soil, move content to the desired location (bin), and rotate anticlockwise to open bucket and unload its content. Participants performed two tasks, and the order of the tasks was randomized among all participants.

Task #1:

Dig soil from the marked area to fill *bin #1* (bin to the left of trench). Accomplish this by using the stylus of Phantom Omni device to control and manipulate the boom/bucket assembly of the simulated excavator. When the bin is full, there will be an audio alert and the content of the bin turns green.

Task #2:

Dig soil from the marked area to fill *bin #2* (bin to the right of trench). Accomplish this by using the stylus of the Phantom Omni device to control and manipulate the boom/bucket assembly of the simulated excavator. When the bin is full, there will be an audio alert and the content of the bin turns green.

3.5.6 Workload Assessment.

Workload is a hypothetical human-centered construct that represent the cost incurred by an operator to achieve a particular level of perception (Sheridan & Simpson, 1979). Therefore, in order to assess the cost in terms of operator fatigue, stress, error, etc.

associated with operating the haptic-controlled excavator for a long period of time, the NASA TLX workload assessment was used. The NASA-TLX is a multidimensional, self-reported assessment technique that provides an estimate of total workload based on six underlying psychological factors associated with task performance. The six underlying psychological factors that contribute to total workload are mental demand, physical demand, temporal demand, performance, effort and frustration level (Hart & Staveland, 1988). Though there are other workload metrics, the NASA TLX was used in this study due to its high validity, applicability, ease of use and popularity in usability research. This was done to provide an understanding of operators' perceived workload, so designers can mitigate their potential impact on system performance.

3.5.7 Procedure.

Participants were briefed on the purpose of the study upon arrival, and then asked to read and sign a consent form. They were briefed on how to complete a computer-based NASA TLX workload assessment after which a pre-test questionnaire was administered to collect demographic information. Participants were informed that their eye movements would be recorded with a remote desktop Tobii® Eye Tracker T60 and that they should maintain a steady head position as much as possible during the test. A short demo of the simulation was given, and participants were given a few minutes to familiarize themselves with the simulator. Questions about the simulator and controls from participants were answered by the experimenter after which actual testing started.

To take the test, participants were seated in front of Tobii® Eye Tracker T60 with their heads about 60cm from the monitor. Participants' head positions were adjusted so that their head was in the middle of the monitor when viewed from behind. Once the appropriate head position is found, participants' eyes were calibrated. This was done by asking the participant to follow the red calibration dot/ball with their eyes as it moved randomly across the screen, briefly stopping at each of the four diagonals and the center of the screen. After calibration, the excavator simulation was initiated and participants were asked to carry out the assigned digging tasks with the Phantom device while their eye movements were recorded with the Tobii® Eye Tracker. The experimental set up for this study is shown in Figure 3.6. Upon completion, participants were thanked, debriefed, and asked to complete the NASA TLX workload assessment and a post-test questionnaire. They were also asked for comments about their experience of using the haptic control excavator interface. Overall, the test took about one hour to complete.

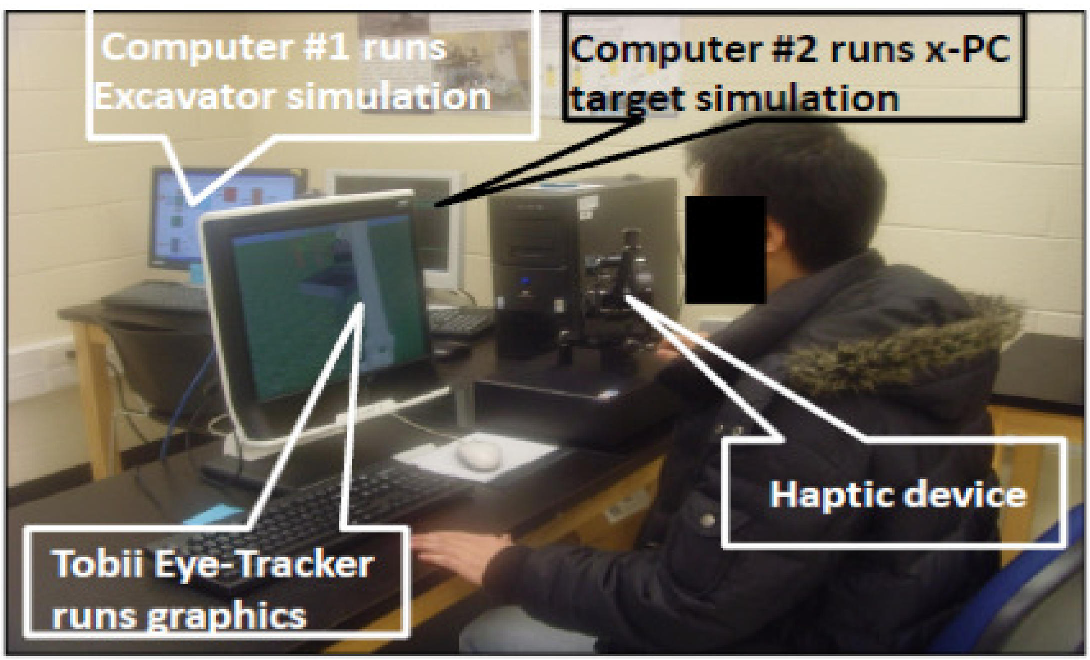

Figure 3.6: Experimental set-up for conflict study

3.6 Results

Performance measures of task completion time (s), number of scoops to fill a bin, and number of scoops dropped outside of bin for both expert and novice operators together with their standard deviations are provided in Table 3.1, while Table 3.2 provides the NASA TLX subjective workload assessment results obtained from participants for each workload metric and total workload.

Table 3.1: Descriptive statistics for task completion time, number of scoops and number of drops

Expertise	Statistic	Performance Measure		
		Completion Time (s)	No of Scoops	No of Drops
Novices	Mean	216.07	7.63	0.85
	Std.dev	67.60	1.36	1.03
Experts	Mean	138.13	6.38	0.25
	Std.dev	6.41	0.25	0.50

Table 3.2: Descriptive statistics for NASA TLX subjective workload assessment for experts and novices

Expertise	Statistics	Workload Metric						
		Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	**Total Workload**
Novices	Mean	12.30	12.13	5.38	6.49	13.11	9.78	59.19
	Std.dev	7.92	9.68	5.84	5.37	8.02	8.83	19.54
Experts	Mean	6.00	5.67	1.92	6.92	11.83	5.50	37.83
	Std.dev	4.69	1.59	2.01	4.22	2.81	5.51	4.96

Several constraints and limitation were encountered in this study that may have influenced the results reported. First, auditory and haptic factors were constant i.e. only one level of haptic feedback was used; similarly, only one auditory alert was used to signal the end of task completion. Second, due to the task domain (i.e. excavation using haptic-controlled excavator), it was difficult to investigate the impact of each modality separately as visual cues were always present. Further, auditory feedback was not used as it did not provide any information because of the lab set up.

3.6.1 Research Question 1. Conflict Detection.

This analysis investigates conflicts between the sensory modalities (auditory, visual and haptic) that may exist in the haptic control excavator interface. In order to probe these conflicts, tasks that depend on auditory, visual and haptic cues were analyzed for expert and novice operators using the dynamic area of interest tool (dynamic AOI) within Tobii® Eye Tracker. The descriptive statistics show that both mean fixation count and mean fixation length were higher for novice operators than they were for expert operators within the area of interest (AOI). Similarly, mean fixation count and mean fixation length outside AOI were higher for novice operators than for expert operators. The results are shown in Tables 3.3 and 3.4. The mean number of fixation count for experts and novices were 190.13 and 281.35 respectively, while mean fixation length for experts and novices were 0.671 and 0.752 seconds respectively on AOI.

Table 3.3: Descriptive statistics for fixation count, fixation length and fixation duration within AOI for experts and novices

Expertise	Statistics	Fixation Count	Fixation Length (s)	Fixation Duration (s)
Novices	Mean	281.35	0.752	0.195
	Std.dev	120.98	0.247	0.416
Experts	Mean	190.13	0.671	0.147
	Std.dev	43.53	0.075	0.068

Table 3.4: Descriptive statistics for fixation count and fixation length outside AOI

Expertise	Statistics	Fixation Count	Fixation Length
Novices	Mean	7.975	0.164
	Std.dev	22.56	0.100
Experts	Mean	1.625	0.095
	Std.dev	1.061	0.048

The mean number of fixation count that fell outside AOI for experts and novices were 1.625 and 7.975 respectively, while the mean fixation length outside AOI were 0.095 seconds and 0.164seconds respectively for expert and novice operators. Figures 3.7*a-d* show graphical representation of fixation count and fixation length for experts and novices respectively within AOI versus those outside of AOI. The notation *(PE R)* in Figures *3.7a* and *3.7c* represents the performance of experts, while *(PR)* in Figures *3.7b* and *3.7d* represents the performance of novices.

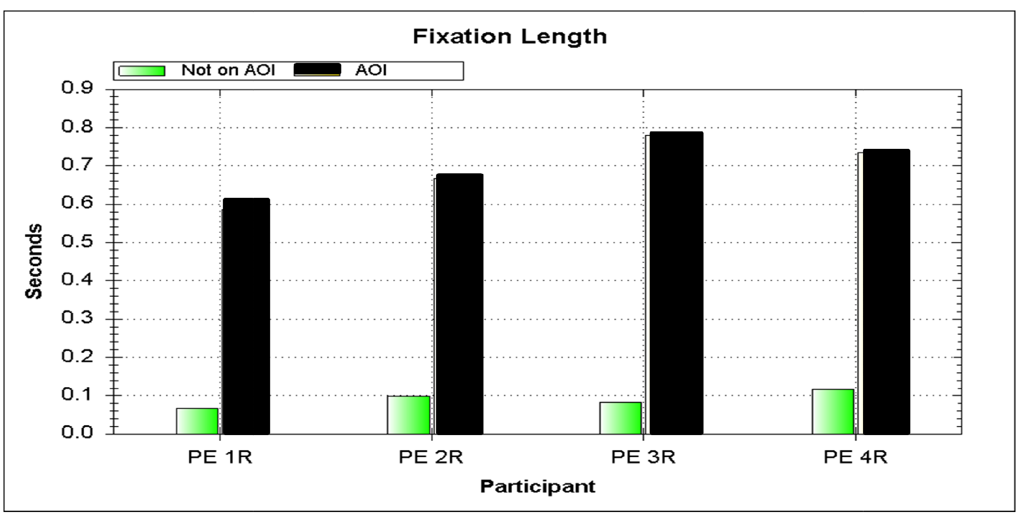

Figure 3.7a: Fixation length within AOI vs. fixation length outside AOI for experts

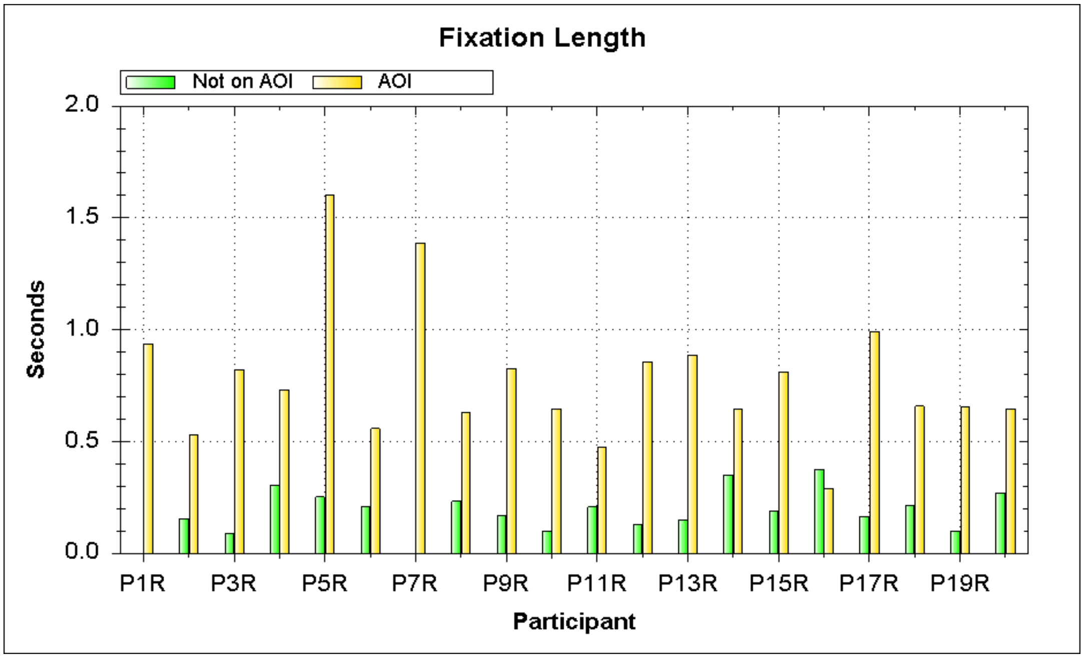

Figure 3.7b: Fixation length within AOI vs. fixation length outside AOI for novices

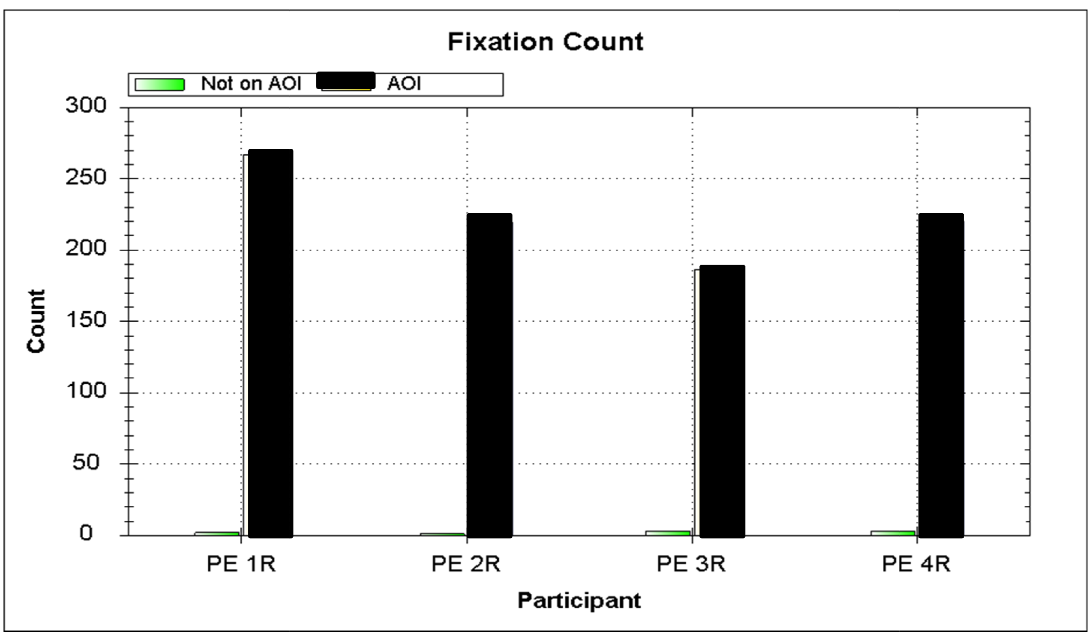

Figure 3.7c: Fixation count within AOI vs. fixation count outside AOI for experts

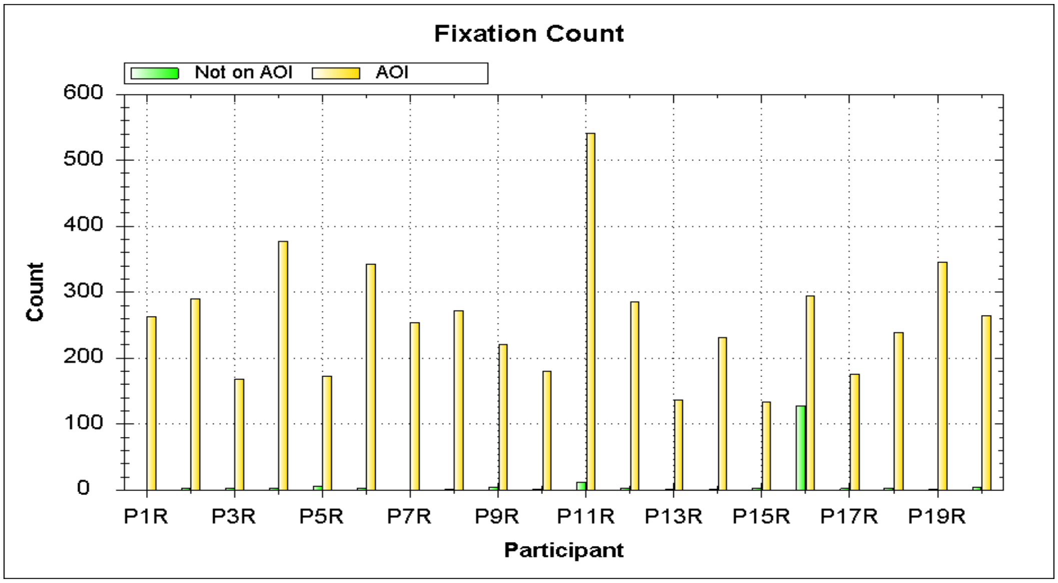

Figure 3.7d: Fixation count within AOI vs. fixation count outside AOI for novices

Prior to using any statistical tool to conduct any statistical analysis, a normality test, test for independence and test for homogeneity of variance (HOV) were performed on each of the data sets and compared to α=0.05 significance level. Normality test using the Shapiro-Wilk's test revealed violation of normality assumption for fixation length *($w=0.9954$* and *$p=0.0009$)*, fixation count *($w=0.8304$* and *$p=0.0001$)*, and fixation duration *($w=0.2885$* and *$p=0.001$).* Normality plots and histograms of the data sets can be seen in Appendix 1. Further, Levene's test for homogeneity of variance showed high variety for fixation length *($F_{1, 46}=4.55$*, and *$p=0.0383$)*, fixation count *($F_{1, 46}=2.52$*, and *$p=0.01192$)*, and fixation duration *($F_{1, 46}=0.47$*, and *$p=0.4958$)*. From the results above, model adequacy was not met, therefore, a non-parametric statistical analysis, the Mann-Whitney-Wilcoxon test was used to analyze the data sets.

Results from the non-parametric Mann-Whitney-Wilcoxon test showed that there was no statistically significant difference in fixation count outside AOI between experts (

$\bar{x}$=8.54) and novices ($\bar{x}$=11.08), *(z=48.0*, and *p=0.9054)*. However, there was a statistically significant difference in fixation count within AOI between experts ($\bar{x}$=13.06) and novices ($\bar{x}$=26.78), *(z=104.50*, and *p=0.018)*. Further, the results showed that within AOI, there was a statistically significant difference in fixation length between experts and novices, *(z=23.00*, and *p=0.0398)*. However, outside AOI, there was no statistically significant difference in fixation length between experts ($\bar{x}$=20.626) and novices ($\bar{x}$=25.275), *(z =165.00*, and *p=0.3988)*.

The higher number of fixation count and fixation length within AOI by expert operators, a difference which the results above show are significant, may be due to the fact that novice operators had harder time keeping their eyes focused in the task area compared to expert operators. In fact, the results show that, novices were nearly twice as likely (3.78 vs. 2.0) to look outside the area of interest while performing the task than experts as seen in Table 3.5 below. This may be due to the interference between the sensory cues that are required for successful execution of the excavation task. The fact that fixation count and fixation length values were higher for novices than for experts may be due to the fact that novices had more difficulty in extracting useful information necessary to execute the task compared to experts.

Table 3.5: Mean number of scan paths outside the area of interest (AOI)

Expertise	Statistics	Mean # of Scan Paths Outside AOI
Novices	Mean	3.78
	Std.dev	1.44
Experts	Mean	2.00
	Std.dev	0.56

The high fixation count also might be an indication that, novices were less efficient in performing the assigned task compared to experts. Thus while experts focused most of their attention within the task area (the screen), novices were unable to focus their full attention on the task area, but alternated between looking at the screen and their hands. The results from this study, therefore, show the existence of possible interference between visual, haptic and auditory cues in the haptic control excavator interface.

3.6.2 Research Question 2. Impact of Conflict on Performance.

This analysis was conducted to investigate whether conflicts between visual, haptic and auditory cues in the haptic control excavator interface have a significant impact on operator performance. By measuring task completion time, number of scoops required to fill up a bin, and number of scoops dropped outside of bin, the performance of experts and novices were compared. Further, error rate (calculated as percentage of number of scoops dropped outside of bin to total number of scoops required to fill the bin) for experts and novices was compared. As can be observed from Table 3.6 below, experts had a mean task completion time of 138.26 seconds with a standard deviation of 6.41seconds while novices completed the task in 216.08 seconds with a standard deviation of 67.60 seconds. Experts filled up bins in 6.375 scoops with a 4% error rate and standard deviation of 0.25, while novices filled up bins in 7.625 scoops with 11.15% error rate and standard deviation of 2.36. Figures 3.8-3.10 show the graphs of mean task completion time, mean number of scoops and error rate for expert and novice operators respectively. The mean number of scoops dropped outside of the bins by experts was 0.25 scoops per bin, while for novices the mean number dropped outside the bin was 0.85

scoops per bin. Prior to conducting statistical analysis using the mean task completion time, mean number of scoops and mean number of drops outside the bin, a normality test, test for independence and test for homogeneity of variance (HOV) were performed on each of the data sets to detect any violations of model adequacy.

Table 3.6: Descriptive statistics for task completion time, number of scoops, number of drops and percentage error

Expertise	Statistics	Completion Time (s)	No of Scoops	No of Drops	Error rate (%)
Novices	Mean	216.07	7.63	0.85	11.15
	Std.dev	67.60	1.36	1.03	9.86
Experts	Mean	138.13	6.38	0.25	4.00
	Std.dev	6.41	0.25	0.50	3.38

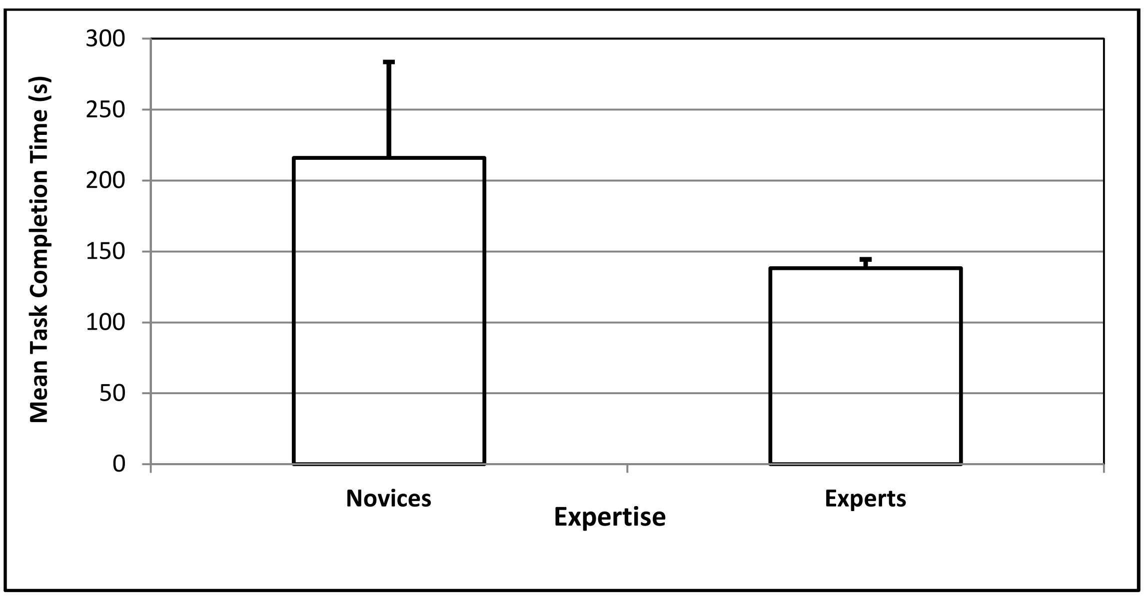

Figure 3.8: Mean task completion time for expert and novice operators

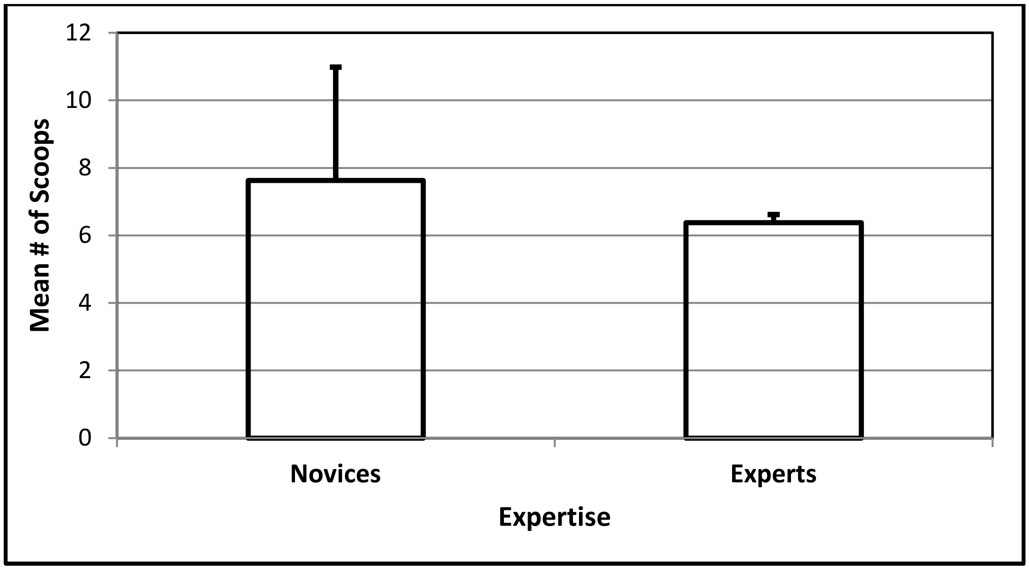

Figure 3.9: Mean number of scoops required to fill a bin by experts and novices

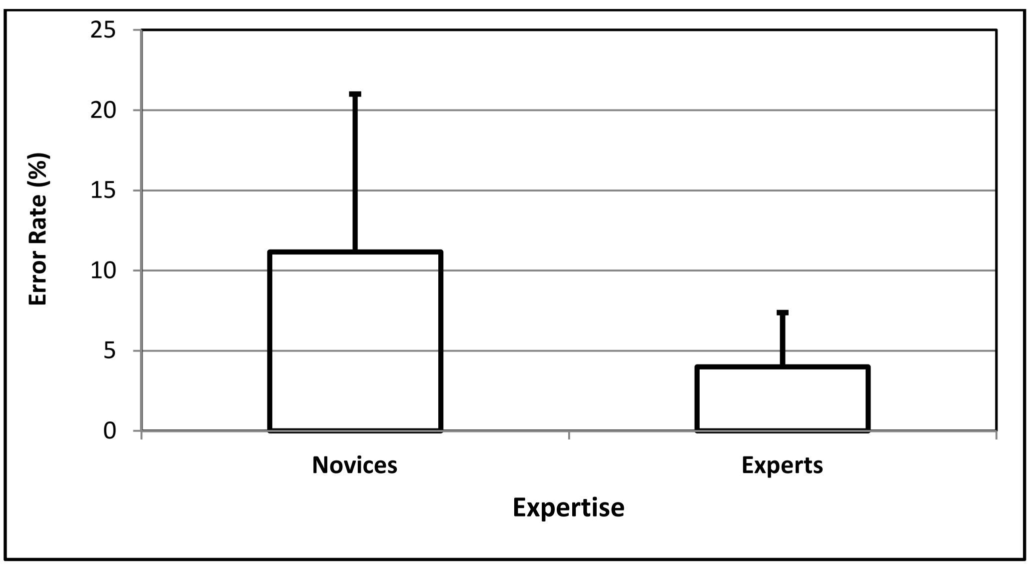

Figure 3.10: Mean error rate for experts and novices

Results of normality test using the Shapiro-Wilk's test revealed violation of normality assumption for mean task completion time *(w=0.8617* and *p=0.0036)*, mean number of scoops to fill a bin *(w=0.8295* and *p=0.0009)*, and mean number of drops outside bin *(w=0.7541* and *p=0.001)*. Further, Levene's test for homogeneity of variance showed *($F_{1, 22}$ =2.63,* and *p=0.1190)*; *($F_{1, 22}$ =3.45,* and *p=0.0768)*; and *($F_{1, 22}$=3.76,* and *p=0.0655)* respectively for mean task completion time, mean number of scoops and mean number of drops outside bin. The analysis showed model adequacy was not met and, therefore, a non-parametric statistical analysis, the Mann-Whitney-Wilcoxon test was used to analyze the data sets. Normality plots and histograms of the data sets can be seen in Appendix 1.

Results from the non-parametric Mann-Whitney-Wilcoxon test showed that there was a statistically significant difference in mean task completion time between experts ($\bar{x}$= 3.25) and novices ($\bar{x}$=14.35), *(z=18,* and *p=0.0047)*. Further, the results showed a statistically significant difference in the mean number of scoops required to fill a bin between experts ($\bar{x}$=5.25) and novices ($\bar{x}$=13.90), *(z=21,* and *p=0.0240)*. However, the results showed that there was no statistically significance in the mean number of drops outside of the bin between experts ($\bar{x}$=9.375) and novices ($\bar{x}$=13.125), *(z=37.50, p=0.6059)*.

The results show that the performance of experts was significantly better than that of novices, in terms of task completion time and the number of scoops required to fill up a bin, however, the performance of experts was not statistically different from novices in terms of the number of scoops dropped outside of the bin. This may be due to the fact that experts had a higher fixation count in the area of interest than novices, which may be attributed to the fact that experts were able to focus their attention in the work area where the actual excavation task took place, while novices wandered in and out of the area of interest. Further, the results also show that, training can be used to greatly improve the performance of novice operators.

3.6.3 Research Question 3. Hand-Eye Coordination.

This analysis was conducted to investigate whether operators had difficulty coordinating their hand-eye movement. To probe whether experts and novices struggled to coordinate their hand-eye movements, the eye-tracking data obtained from the study was analyzed. Fixation count within and outside the AOI, fixation length within and outside AOI, as well as scan paths were analyzed. As shown in Table 3.7, within AOI, experts had lower mean fixation count than novices (190.125 vs. 281.35) and mean lower fixation lengths than novices (0.671seconds vs. 0.752 seconds). A non-parametric Mann-Whitney-Wilcoxon test was performed since data set violated normality test as discussed in Section 3.6.1.

Table 3.7: Mean fixation count and fixation length for experts and novices within and outside AOI

		Experts	**Novices**	**z-value**	***p***
Within AOI	Mean Fixation Count	190.125	281.35	104.50	0.0118
	Mean Fixation Length (s)	0.671	0.752	23.0	0.0398
Outside AOI	Mean Fixation Count	1.625	7.975	48.00	0.9045
	Mean Fixation Length (s)	0.095	0.164	16.50	0.3955

Results from the Mann-Whitney-Wilcoxon test showed that within AOI, there was a statistically significant difference between experts and novices ($z=104.50$, and $p=0.0118$) in mean fixation count, similarly, there was statistically significant difference between experts and novices in mean fixation length ($z=23.0$, and $p=0.0398$). Outside AOI, there was no statistically significant difference in mean fixation count between experts and novices *(z=48.0,* and *p=0.9054)*, similarly, there was no statistically significant difference between experts and novices in mean fixation length *(z=16.5*, and *p=0.3988).*

Further, to gain an understanding of operators' mental processes as they carried out the excavation task, the gaze plots and scan path data obtained using eye-tracking were analyzed. Since what the human eye looks at usually reflects what goes on mentally, the gaze plot data was used to gauge operators' mental processes. The gaze plots for expert and novice operators are shown in Figures 3.11 and 3.12 respectively.

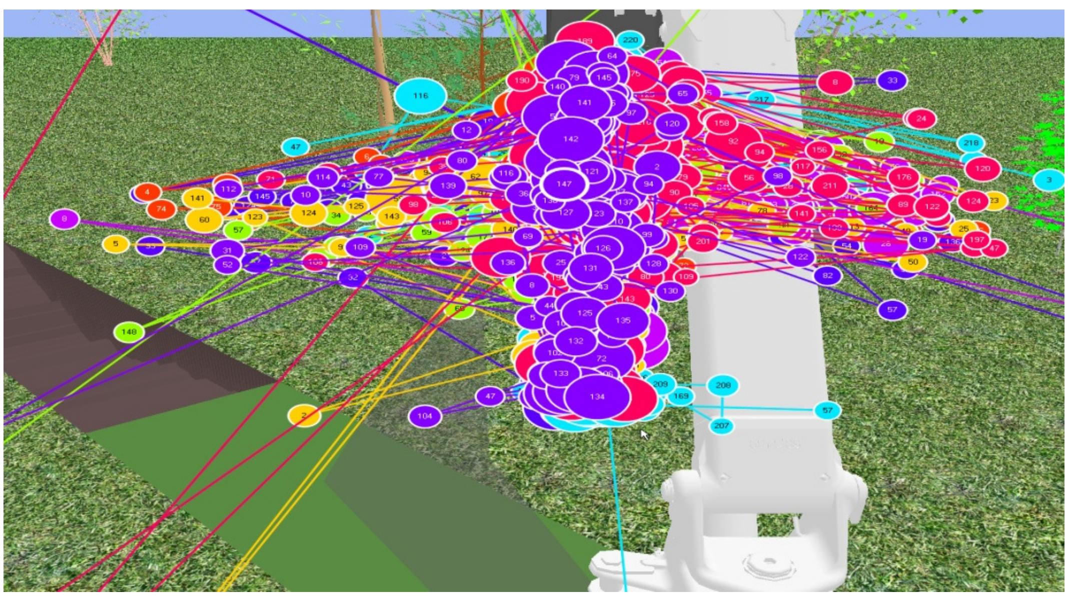

Figure 3.11: Gaze plot for expert operators

From the gaze plot data, it was observed that experts' attention were focused on the environment where the task was performed, this is demonstrated by the fact that, most gaze lines of expert operators were within the work area as seen in Figure 3.11. Novice operators on the other hand, were unable to fully focus or limit their eye movements to the work area as demonstrated by numerous gaze lines that go off the screen as seen in Figure 3.12.

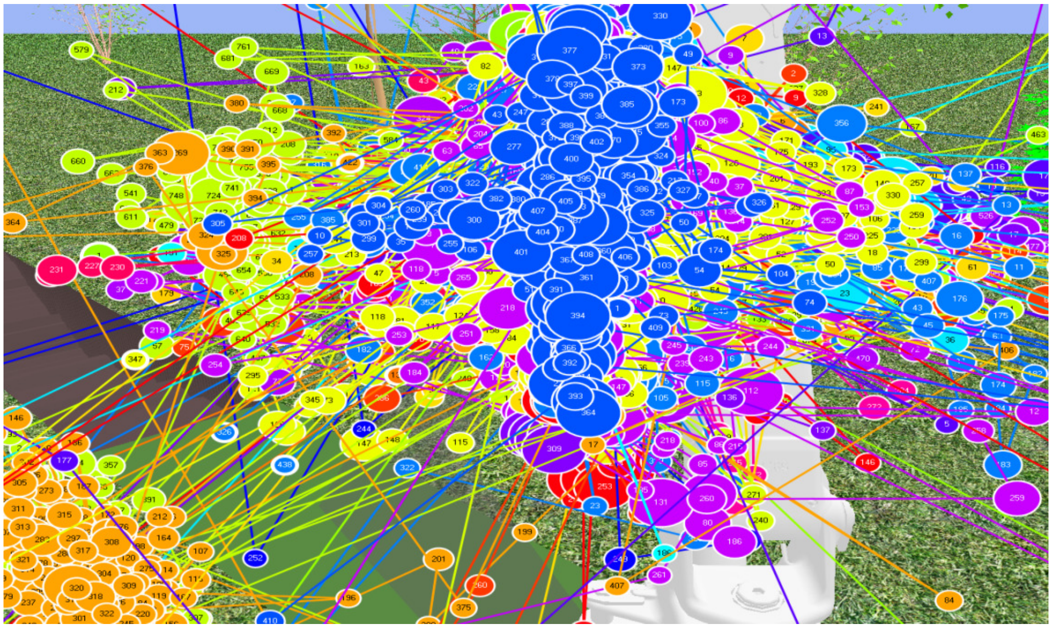

Figure 3.12: Gaze plot for novice operators

The off screen gaze lines are indication of novices attempting to look at their hands as they performed the excavation task with the phantom device. This may be due to the fact that, novices struggle to keep their eyes focused on the screen where the actual excavation task takes place, but rather keep their eyes from looking on the screen to looking at their hands. A situation similar to an experienced driver's ability to accelerate and brake while driving vehicle without having to look at the accelerator or brake pedals, an involuntary action. On the other hand, an inexperienced driver might be tempted to look at the accelerator or brake pedal in order to move or stop a vehicle, a voluntary action.

In summary, the results show that operators had difficulty coordinating their hand-eye movement while operating the haptic-controlled excavator. Further, the results show that, novice operators had more difficulty coordinating their hand-eye movement than did expert operators. This may be due to the fact that experts were able to retrieve information from memory to help them accomplish the task while novices did not.

3.7 Chapter Summary

In order to design an effective, intuitive and easy to use interface, it is important that the complimentary sensing cues are integrated in a way that capitalizes on the strengths of each mode in order to overcome the weakness in each other. To design a robust and easy to use haptic-controlled excavator interface, it is important that issues of conflict and interference between the multiple sensing cues used in the design are well understood. To accomplish this, an empirical study was conducted to assess whether conflict exists between visual, haptic and auditory cues that are necessary for the smooth operation of the haptic control excavator interface.

The goal of the empirical study was to identify if there were conflict between visual, haptic and auditory cues in the haptic interface, and whether these conflicts had an impact on the performance of the operator. Results from the empirical study show that conflicts do exist between the visual and haptic modalities in the haptic control excavator interface, and this interference does impact the operation of the haptic control excavator. From the results, performance of novice operators was impacted more by the interference between the sensory cues than the performance of expert operators. Finally, results from the empirical study show that novice operators had a harder time coordinating their hand-eye movement than expert operators. Overall, the results from the empirical study provided an understanding of the interference between the sensory modalities and their effects on operator performance.

CHAPTER 4

HAPTICS IN FLUID POWER SYSTEMS

4.1 Background

Touch is the fundamental attribute of interpersonal communication that makes human-human interaction so natural, intuitive and rich in information. Whether a greeting handshake, an encouraging pat on the back, or a comforting hug, physical contact is a basic means through which people achieve a sense of connection, indicate intention, and express emotion (Brave & Dahley, 1997). Touch is even more valuable in close personal relationships, such as family and friends, where it is often used to express affection. Haptic, or touch is omnipresent of everyday human activity and provide continual and essential source of information during the performance of virtually any physical activity ranging from reading a book, where we almost subconsciously hold and turn the pages, to participating in a sport, where proficiency in haptic interaction is highly prized and honed to near perfection (Oakley et al., 2003). Further, haptic feedback is of critical importance whenever humans interact with objects in our environment, either by picking objects or interacting and manipulating objects in some way, humans instinctively rely on the inherent haptic cues and feedback received from these interactions to inform us about the properties of the object such as its texture, shape, weight, hardness, stiffness etc.

Though haptic or touch sense is used by humans to interact with environment, compared to visual and auditory senses, understanding of human haptics, which includes the sensory and motor systems of the hand, is very limited. One of the reasons for this lack of understanding of haptics is the apparent difficulty to experimentally present control haptic stimuli, mainly due to the bidirectional nature of haptics (haptics can simultaneously be perceived and act upon the environment).

Haptic interfaces are devices which are composed of mechanical components in physical contact with the human body (hand) for the purpose of exchanging information with the human nervous system, therefore, when performing tasks with a haptic interface, the human user conveys desired motor actions by physically manipulating the interface, which, in turn, displays tactual sensory information to the user by appropriately stimulating his or her tactile and kinesthetic sensory systems (Biggs & Srinivasan, 2002; Srinivasan, 1995).

Haptic interfaces according to Srinivasan (1995) can thus, be viewed as having two basic functions: first to measure the positions and contact forces (and time derivatives) of the user's hand and/or other body parts, and second, to display contact forces and positions and/or their spatial and temporal distributions to the user. He further argues that, among these position (kinematic) and contact force variables, the choice of which variables are considered motor action variables (i.e. inputs to the computer) and which ones are considered sensory display variables (i.e. inputs to the human) depends on the design of hardware and software, as well as the tasks the interface is designed for, and that most current force reflecting haptic interfaces sense position of their end-effector and display forces to the human user. In many respects, haptic interface device is analogous to mouse, except that the mouse is passive and cannot communicate with the user, while the haptic device can provide force feedback to the user through haptic rendering.

4.2 Biomechanics of Touch

The psychophysics of touch or the study of how humans perceive touch is vital to the design of efficient and responsive haptic interfaces. When humans touch objects, tactile information is provided by the spatio-temporal distribution of the mechanical loads on the skin at the point of contact (Biggs & Srinivasan, 2002). Primarily, humans perceive touch through one of two perceptual systems, the cutaneous/ tactile system or the kinaesthetic system. While the cutaneous or tactile system, refers to information sensed through the medium of skin, and encompasses such disparate sensations as texture, temperature and pain, the kinaesthetic system, refers to stimuli originating from an intimate knowledge of the internal state of the body (Burdea & Brooks, 1996).

The ability of the tactile sense to perceive vibrations enables humans to distinguish among a wide variety of textural information while kinaesthetic sense enables humans to be aware of the positions of the limbs and the forces exerted by the muscles (Oakley, 2003). The tactile and kinaesthetic senses yield distinct but complementary information such that the tactile sense provides information about the fine grained details of an object, such as its texture, while the kinaesthetic sense informs us about the larger scale details of an object, such as its shape, weight, hardness, stiffness etc (Oakley, 2003). The finger pad which is tactilely sensitive than any other part of the human body, can detect the location of a point to within 0.15mm, detect two points that are approximately 2mm apart and detect a dot 2 microns high on a smooth surface (Johnson & Phillips, 1981).

The finger pads of humans are made up of complex sensory structure which contains receptors (proprioceptors) in both the skin and the underlining tissues (Bem, 2011; Strauss, 1999). These receptors carry signals to the brain. Whenever humans touch an object, contact is made between the finger pads and the surface of the object. As the hand reaches the object, it adjusts to the shape of the object and generates a unique set of data points that describe joint angles, muscle length and tension. The information/data collected by the receptors is sent to the brain where it is processed allowing the brain to understand the subtle tactile details (smoothness, coarseness, hardness, etc) about the object. Similarly, changes in muscle tension are processed to provide kinesthetic information (size, shape, position, etc.) of the object. The tactile and/or kinesthetic feedback that the human receives is referred to as haptic force feedback. Haptic perception incorporates both touch stimuli from the skin and kinaesthetic stimuli from the position and movement of joints and muscles. Unlike visual and auditory modalities, haptic design is nearly always a multimodal design: haptic is generally used in conjunction with other sensory modalities, usually to reinforce same tasks or to handle different tasks performed at the same time. In a multimodal environment, where the user's primary attention as well as visual resources and possibly hands are engaged in other tasks, touch cues presented to the skin can be used to notify the user of events and to create relatively unintrusive, ambient background awareness (Hayward & MacLean, 2007; MacLean & Hayward, 2008).

4.3 Applications of Haptics

It is not difficult to identify ways in which haptics could be applied to aid humans. With advances in computing technology and the need for better and intuitive interaction between humans and machines, a number of researchers and universities are experimenting with haptics. Among the numerous applications as outlined by Hayward et al. (2004) include:

1) Force-reflecting input devices for use with graphical user interface augmentation, where haptic cues are provided to enhance existing graphical interfaces to increase efficiency, speed and reduce fatigue. For example, haptic cues have been integrated into visualization tasks to allow researchers to interact haptically with data and to present mathematical data to the visually impaired.
2) In teleoperation or telerobotics, a human operator controls the movement of a remote robot by relying on haptic feedback received from the remote robot. This is usually accomplished through remote manipulation of some distant robotic device, referred to as slave, by the manipulation of a local robotic device, called master. In a master-slave teleoperation, user manipulations of the master device are reflected in the slave device, and user receives haptic feedback from the environment through the slave-master arrangement. The goal is to provide a representation of the objects that the slave physically encounters in the remote environment to the human operator through the sense of touch. The result of this representation is that the operator feels as though he/she is located in the remote environment. Telerobotics and remote manipulations have been traditionally

applied to perform complex, human controlled, physical manipulations in environments that are inhospitable by humans such as an environment contaminated by radioactive material, deep under water environment or space (Cooper, 1998). Remote manipulation has also been widely used to perform dangerous activities such as bomb disposal as well as for search and rescue missions in disaster zones that may be impossible humans to operate in.

3) The use of force feedback in interacting with virtual environments has proven very successful for applications in entertainment industry. Video game makers have adopted passive haptics to take advantage of vibrating joysticks, controllers and steering wheels to reinforce on-screen activity. The use of force feedback in video games provides users an increased sense of involvement in the simulated environment and heightened sense of realism, all of which lead to improved user experience. Available literature indeed, suggests that the presence of force feedback in video games results in increased feelings of immersion in the virtual world (Oakley, 2003).
4) Vehicle operation and control room operation to alleviate visual load in stressful and fast-paced environments.
5) Medical robotics allows surgeons to reach organs and tissues that will otherwise be difficult to reach with minimal invasion, and training through simulation. Traditionally, surgical procedures training require skilled physical tasks, in which practitioners rely heavily on their sense of touch (Burdea & Brooks, 1996), and involve long periods of apprenticeship during which substantial number of

operations are first observed and then performed. This process posses several challenges, first, it posses danger to patients, second, it may prolong operation time and increase cost, and third, there is a lack of suitable patients for training which may lead to omission or reduction of training in a particular procedure. Hence, virtual reality simulations with haptic feedback could be used to augment traditional training techniques by providing simulated procedures through which students can practice and learn without the need for real patients. However, in order for these simulations to provide valuable training, they must achieve realism which often entails simulating the feel of a variety of medical implements as they grasp or cut and also simulate the behavior of deformable organic surfaces (Oakley, 2003).

6) Force feedback is used as a physical rehabilitation tool to train stroke patients, and improve working conditions for visually impaired. Most stroke survivors have a natural tendency to overuse their less-affected arm or leg in performing activities of daily living. To help overcome the overuse of the less-affected arm, and to increase the productive use of the impaired arm, active force feedback could be embedded into a meaningful driving simulation environment to create a robot-assisted therapy device which then motivates patients to engage the impaired arm and aid rehabilitation (Johnson et al., 2005; Popescu et al., 1999). Another area where force feedback presents huge potential is in education and training such as surgical training, dangerous systems or systems with limited availability such as surgical patients could be simulated using haptics (Hayward et al., 2004). For

example, the mechanical engineering department at Rice University has adopted a haptic paddle interface for the teaching of dynamic systems course to undergraduate students. Preliminary assessment showed that using the haptic interface improved students learning of dynamic systems concepts when compared to traditional teaching methods (Bowen & Marcia, 2006). Other application areas include engineering such as computer aided design, arts and graphic design for creation of animation, editing sounds and images as well as manufacturing to assist assembly design and reduce prototyping.

4.4 Haptic, Visual and Auditory Modalities in Time and Space

In order to develop a multimodal human-machine interface for application in fluid power systems, it is important to understand the synergistic relationships between haptic, visual and auditory modalities and how they exist in time and space. Gaver (1989) developed a model of the existence of sound and vision in time and space shown in Table 4.1. Gaver (1989) used this model to argue that humans use visual and auditory information synergistically in daily activities, not only to increase the bandwidth of available information, but also because the visual and auditory information complement each other. For example, during casual conversation, one may read the lips of the speaker in order to fill in missing words. Based on Gaver (1989), a two-dimensional synergistic framework for the existence of touch, vision and sound in time and space is proposed as seen in Figure 4.1. This framework expands Gaver's original model to include touch/haptic modality.

Table 4.1: Synergistic modes of vision & sound in time and space [Courtesy of Gaver, 1989]

	TIME	SPACE
SOUND	**Sound exists *in time*** (a) **Good for conveying changing events** (b) **Available only for a limited time**	**Sound exists *over space*** (a) **Receiver does not need to face source** (b) **A limited number of messages can be conveyed simultaneously**
VISION	**Visual objects exists *over time*** (a) **Good for display of static objects** (b) **Can be accessed repetitively over time**	**Visual objects exists *in space*** (a) **Receiver must face source** (b) **Messages can be spatially distributed**

The goal is to understand the synergistic relationships between vision, sound and touch in time and space, and the impact of this synergy on multimodal human-machine interface. Only haptic, visual and auditory modalities are considered in this framework because the other two human senses (taste and smell) are currently not well developed/natured technologically to be used efficiently and effectively in human-machine system interaction. The key features of the expanded framework (i.e. the existence of haptics in time and over space, the existence of sound in time and over space, and the existence of vision in space and over time) are described below. In general, the existence of sound and haptic in time and space share some similarities, while the existence of vision in time and space is not shared with any other modality.

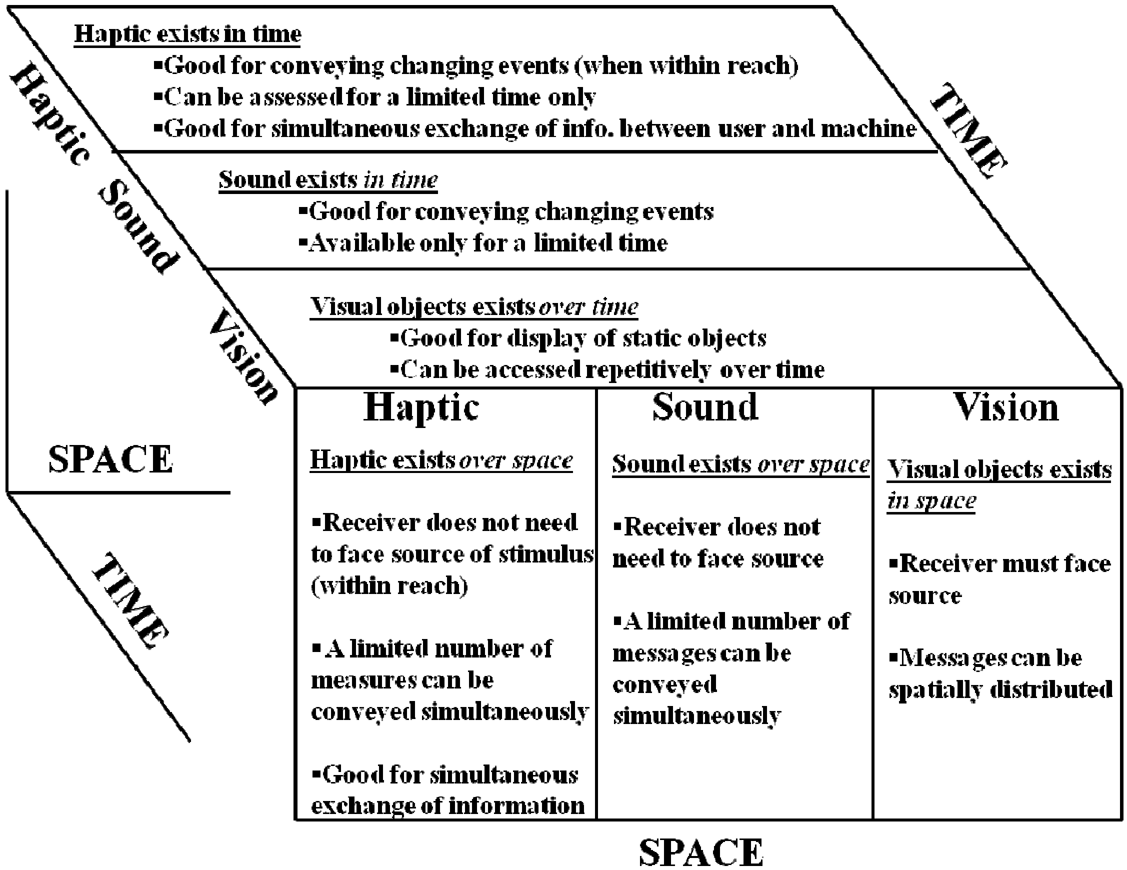

Figure 4.1: Synergistic framework of vision, sound and touch in time and space

4.4.1 Haptic Exists in Time and Over Space.

The haptic modality shares many similarities with auditory modality in its existence in time and space. The haptic modality exists in time because like sound, it is inherently ephemeral in nature, i.e. touch is short-lived with a beginning and an end. For example, we can touch an object to feel whether it is hot or cold, hard or soft, smooth or rough, heavy or light; we can also touch to express emotions such as touching a friend to express affection or love etc. In general, humans depend on the sense of touch to determine object hardness, weight, inertia, contact geometry, smoothness, slippage and temperature. The experience of touch in nature is also brief in most cases that it could be experienced. As a result, the sense of touch is most suitable for conveying information about changing events to a single user, especially in situations where the target recipient is an individual in which case sound may be unsuitable because sound may alert others around too. For example, we may set our cell phones to vibrate to alert us of incoming calls when we don't want distract the attention of others around us. When driving on the highway, if a driver gets distracted and steers towards the side of the road, the grooves on the sides of the road vibrates the vehicle and alerts the driver to steers back onto the road.

Unlike vision, touch exists over space, i.e. a user does not need to face the source of the stimuli to experience the sense of touch. We can sense when a friend touches us on the back or shoulder, thus, like sound touch can convey information to users irrespective of their orientation, however, for the purpose of humans interacting with computers, the number of messages that could be presented to a user simultaneously through touch sensation may be limited.

Another obvious advantage of the touch modality in human computer interaction is its unparalleled ability to provide for simultaneous and bi-directional information exchange between a user and a machine/computer. Touch, thus, allows both users and computers to experience simultaneous exchange of information through the use of haptic devices. The result is that users '*feel*' engaged in the performing task; however, an obvious disadvantage of the touch modality is that it has limited reach zone and usually used in conjunction with visual or auditory mode in the design of human computer interfaces.

4.4.2 Sound Exists in Time and Over Space.

The existence of sound in time and space very much resembles that of touch sense. It is short-lived, with a start and finish points that could be experienced. As a result, sound is often used to convey information about changing events (Gaver, 1989). Like the sense of touch, a user does not need to faces the source of the stimuli in order to hear a sound. Sound can be heard from all directions irrespective of the user's orientation; therefore, sound is often used to convey warning alerts to users. For example, sound is often used in alarm systems though other modalities such as visual alerts may be used as redundant cues.

Another reason that accounts for the effectiveness of auditory modality in attracting users' attention is due to the fact that it is able to provide information beyond the reach of either visual or haptic modality. Sound has the unique ability to provide information from all the directions to users, as a result, it is possible for people to listen and hear actions even when they cannot see the source. For example, it is possible to hear noise coming from next office or hear revving of a car outside the home even if we cannot see them.

4.4.3 Vision Exists in Space and Over Time.

Unlike sound and touch, vision exists in space, i.e. the user must face the source of the stimuli in order to see. Thus, in order to perceive visual information from the environment, one has to look in the direction of the visual stimuli. It is not possible to perceive visual stimuli if the user is not looking in its direction (a user cannot see from behind). On the other hand, many visual items can be displayed simultaneously provided they are located in the same direction. Vision is, therefore, suitable for displaying static information which does not change but remain stable over time. For example, a user can monitor a panel that has several gauges, or monitor two or three computer screens at the same time etc. Also, unlike sound or touch, visual stimuli can be accessed repeatedly over time. In the design of multimodal operator-excavator interface for haptic-controlled excavator interface, visual, auditory and haptic modalities will be used to complement each other and strengthen the weakness in each other to produce a synergistic blend that is more efficient and effective than could be achieved from each modality working alone.

4.5 Haptic-Controlled Excavator Interface Study

As described in the preceding session, haptics has great potential if properly integrated into the haptic control excavator interface. First, the incorporation of haptics into the user interface of excavator will provide a simultaneous exchange of information between the operator and the excavator, thus, enabling the operator to experience an "immersed" interaction in the environment in which the task is being performed. Thus, in combination with visual display, haptic interface can be used to train operators to better perform digging tasks that require hand-eye coordination, and provide valuable help to novice operators to improve their task performance.

Further, since human cognitive processes and perception build largely upon multimodality, a proper combination of haptic, visual and auditory modalities will result in a flow of information on several parallel channels which has been shown to enhance effectiveness of interaction(Krapichler et al., 1999). By making use of the haptic-controlled interface instead of the traditional levers and pedals, the excavator operators will be freed from solving the inverse kinematic relationship, and result in a more efficient and effective task performance and shorten training time for novice operators (Kontz & Book, 2007).

In traditional excavator operation, operators rely only on visual and auditory information to accomplish the task of excavation, therefore, by incorporating the haptic modality into the new design, a third modality 'haptic feedback' is introduced with the expectation that this extra modality will help reduce the load placed on the visual system and lead to improve operator performance. For example, haptic feedback may help alert operator to the presence of buried unusual/unknown obstacles or objects that may be encountered in the work environment, and thus, help the operator avoid them. It is not uncommon to find excavator operators accidentally causing damage to underground utility lines (water, gas, electric power lines etc) in construction sites primarily due to their inability to see the presence of these lines before hand. With haptic control interface, it is expected that, when the bucket of the excavator encounters such an obstacle, force feedback will be sent to the operator to alerts him/her to the presence of such an obstacle, and enable the operator to perform the task in much more safe and efficient manner. In addition, humans have a natural tendency to interact multimodally with the environment, therefore, a multimodal haptic-controlled interface will be more intuitive, easy to learn and use, and can reduce operator's mental workload and stress level resulting in improved situation awareness, better judgment, and decision making.

Although the haptic interface promises reduced mental workload and improved operator performance over the traditional lever/pedal interface, its use as a control interface for the excavator has not been fully explored because the technology is still being developed. Currently, the concept of haptic-controlled excavator interface is under development at Georgia Institute of Technology. The haptic input device is PHANToM 1.0 originally designed by Salisbury (1995) and subsequently commercialized by SensAble Technologies. In order that the potential benefits of haptics in the excavator interface be realized, it is important to understand the basic biomechanical, sensorimotor, and cognitive abilities of the human haptic system, in order to properly determine the design specifications of the hardware and software of haptic interfaces. The following section describes a pilot study and an empirical study that were conducted to determine the appropriate force feedback values necessary in the haptic control excavator interface for best operator performance.

4.6 Pilot Study

The first challenge of the empirical study was to identify low, medium and high force feedback range values to be used in the empirical study. Since the author was not aware of any experiment that had grouped force feedback values used in a haptic control excavator into low, medium and high, a pilot study was conducted to group the force feedback values into low, medium and high.

4.6.1 Procedure.

To do this, the author first conducted several trials using different force feedback values to assess their impact on performance of the operator. In these initial trials, task completion time was used as the measure/metric for assessing operator performance. This metric was used due to time constraints and also due to the fact that it is the most important metric in the assessment of operator performance. Based on several trials, the author's knowledge of the haptic-controlled excavator, and consultation with subject matter experts (SMEs), 10 possible ranges of force feedback values were identified. This was done by changing the force feedback parameters in the *MatLab* code that ran the excavator simulation and monitoring the effect of the change on operator performance. The initial ranges of force feedback values identified along with "0" (no force feedback) are shown in Table 4.2 below.

Table 4.2: Initial range of force feedback values identified by author

Group	1	2	3	4	5	6	7	8	9	10
Range	0.01-0.05	0.05-0.1	0.1-0.2	0.3-0.5	0.6-0.7	0.8-0.9	1.0-1.2	1.2-1.5	1.5-1.7	1.8-2.0

After the initial ranges of force feedback values have been identified, five volunteers were recruited to participate in a pilot study to investigate how these different force feedback values affected operator performance. To conduct the pilot study, each of the five volunteers were first briefed on the purpose of the study, and then asked to sign a consent form. Volunteers were seated in front of the computer that ran the excavator simulation, and a 10-minutes trial demo was performed to familiarize them with the haptic control excavator. After volunteers became familiar with how to manipulate the haptic control excavator, all their questions were answered by the experimenter and actual testing began. A schematic equipment setup is shown in Figure 3.7 in Chapter 3; the only difference being that unlike the previous study, Tobii® Eye Tracker was not used in the current study.

For each range of force feedback values, volunteers were asked to load one of the bins (bin #1/left side bin) located the haptic control excavator work area by using the stylus of the Phantom Premium 1.5 device to control and manipulate the simulated excavator. The order of loading was randomized among volunteers to eliminate learning effect. Only the left side bin (bin #1) was used for this pilot study because asking volunteers to load both bins (bin #1 and bin #2) would have doubled the amount of time required to complete the experiment. Further, a prior study conducted by the author, Osafo-Yeboah et al. (2010) showed that, task completion time for bin #1 and bin # 2 were highly correlated. Results obtained from the pilot study are discussed in the next Section.

4.6.2 Results from Pilot Study.

The descriptive statistics obtained from the pilot study are summarized in Table 4.3 below. From the descriptive statistics, it was observed that task completion time improved as the force feedback values were gradually increased. This improvement in task completion time continued to a point, and then began to decline as the force feedback values got higher.

Table 4.3: Force feedback range with corresponding mean task completion time and standard deviation

Force Feedback Range (N)	0.01-0.05	0.06-0.1	0.11-0.2	0.3-0.5	0.6-0.7	0.8-1.0	1.10-1.2	1.2-1.5	1.5-1.7	1.8-2.0
Mean Task Completion Time (s)	174.4	171.1	170.8	162.8	151.5	148.3	150.1	167.0	178.4	174.3
Std.dev	46.04	35.12	50.89	51.54	46.88	46.60	52.71	52.12	62.22	59.43

Based on this result, a low, medium and high force feedback range values were classified by combining the force feedback range values with similar task completion times. The resulting classification is shown in Table 4.4.

Table 4.4: Classification of force feedback range values

Force Feedback Range (N)	Classification	Mean Task Completion time (s)
0.0-0.2	low	172.1
0.2-1.2	medium	153.178
> 1.2	high	173.24

Prior to using Analysis of Variance (ANOVA) to analyze results, a residual plot and normality check was performed on data set to ensure no obvious violation. The normality check on task completion time showed no violation of normality using Shapiro-Wilk's test statistic *(*$w = 0.9276$ and $p=0.4267$*)*. Further, no obvious violation of independence and randomness was observed from the residual plot.

An ANOVA was conducted to compare the mean task completion times for the low, medium and high force feedback range values. The ANOVA analysis at $\alpha = 0.05$ showed that there was a statistically significant difference in mean task completion time between low, medium and high force feedback classifications ($F_{2,9} = 10.21$, and $p=0.0084$). Further, a post-hoc Turkey test showed that the mean task completion time for medium classification was different from that of low and high classifications, however, the mean task completion time for the low and high classifications were not different. This means that there is little benefit to operators when force feedback is low, however, as force feedback is increased; performance improves resulting in lower task completion time. At high force feedback, the forces in the haptic devise begin to interfere with task, resulting in reduced task completion time. The mean task completion time for each classification is shown in Figure 4.2. These force feedback range values *Low*, *Medium* and *High* were subsequently used in the empirical study to investigate the range of force feedback that produced best operator performance.

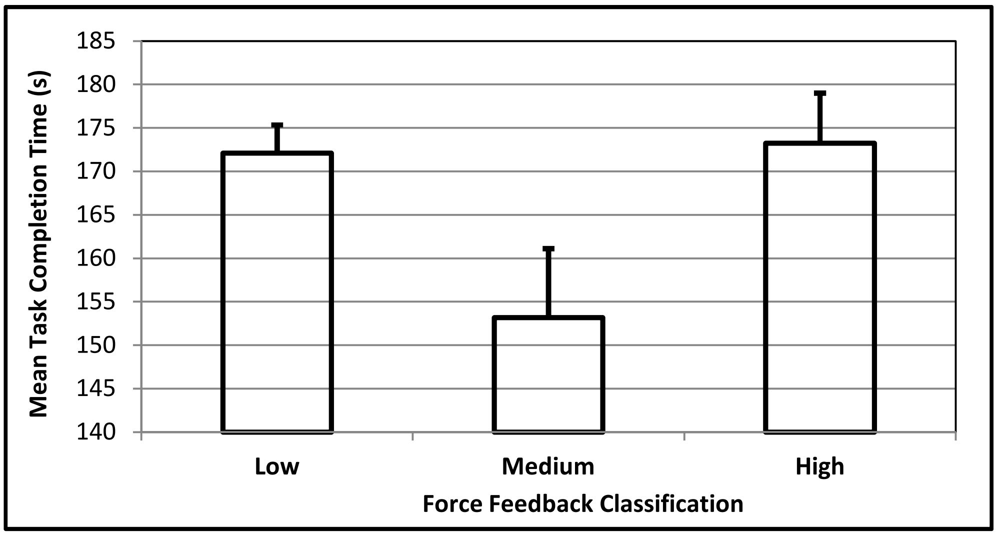

Figure 4.2: Mean task completion time for low, medium and high force feedback

4.7 Methodology

This section describes an empirical study conducted to identify the range of force feedback values that yield best operator performance.

4.7.1 Research Questions.

This experimental study seeks to answer the following research questions.

(1) Does different force feedback levels affect operator performance?

(2) What is the optimal range of force feedback values that yield best operator performance?

4.7.2 Participants.

Twenty students ages 19 to 46 years (mean age = 25.95 and standard deviation =7.06) were recruited from the North Carolina Agricultural & Technical State University to participate in this empirical study. Participants were made up of 12 males and 8 females and consisted of both graduate and undergraduate students. Each participant received a $20 gift certificate for their time.

As described in Section 3.5.2, the sample size for this study was determined by assuming a power, p of 0.8 for the test. The formulation $p = (ES)*\alpha*\sqrt{n/\sigma}$, where p is the power, ES is the effect size, α is the significance level, n is the sample size, and σ is the standard deviation was used. From the pilot study, the effect size (ES) was estimated to be 2, and σ was estimated at 25.29 seconds (0.4215 minutes) from a previous study by the author (Osafo-Yeboah et al., 2010). With a significance level, α of 0.05, sample size was calculated to be 27. However, due to monetary and time constraints, only twenty participants were recruited for this study, and the power was calculated as 0.688.

4.7.3 Experimental Design.

A within-subject design was used in this experiment. The independent variable investigated in this experiment was the range of force feedback with 4 levels (*no force feedback, low force feedback, medium force feedback and high force feedback*). The dependent variables are

(i) task completion time

(ii) number of scoops required to fill a bin

(iii) number of scoops dropped outside of bin, and

(iv) accuracy rate (percentage of the mean number of scoops dropped outside of bin to the mean number of scoops required to fill a bin)

To remove learning effect/carry-over effect, the order of the test was completely randomized. One half of the participants started the test with the no force feedback condition and finished with the high force feedback condition, while the other half started with the high force feedback condition and finished with the no force feedback condition.

4.7.4 Equipment.

The equipment for this empirical study was similar to the setup used in the study described in Chapter 3. It consisted of three Gateway computers and a Phantom Omni 5.3 Haptic device. Computer #1 interfaced with the Phantom Omni and ran the excavator dynamics simulation, computer #2 ran the xPC-target simulation, and computer #3 ran the excavator simulation graphics. All three computers were connected via a local network. The Phantom Omni device sat next to the excavator simulation graphics computer (computer #3) on the right hand side of participants and had 6 degrees of freedom in total: up-down, left-right, front-back, and a rotating stylus with 3 degrees of freedom. The schematic layout of the equipment setup is shown in Figure 4.3.

4.7.5 Procedure.

Participants were first briefed on the purpose of the study upon arrival, and asked to read and sign a consent form. A pre-test questionnaire was administered to collect demographic information. A short demo of the simulation was given, after which participants were given about 15 minutes to try out and familiarize themselves with the simulator.

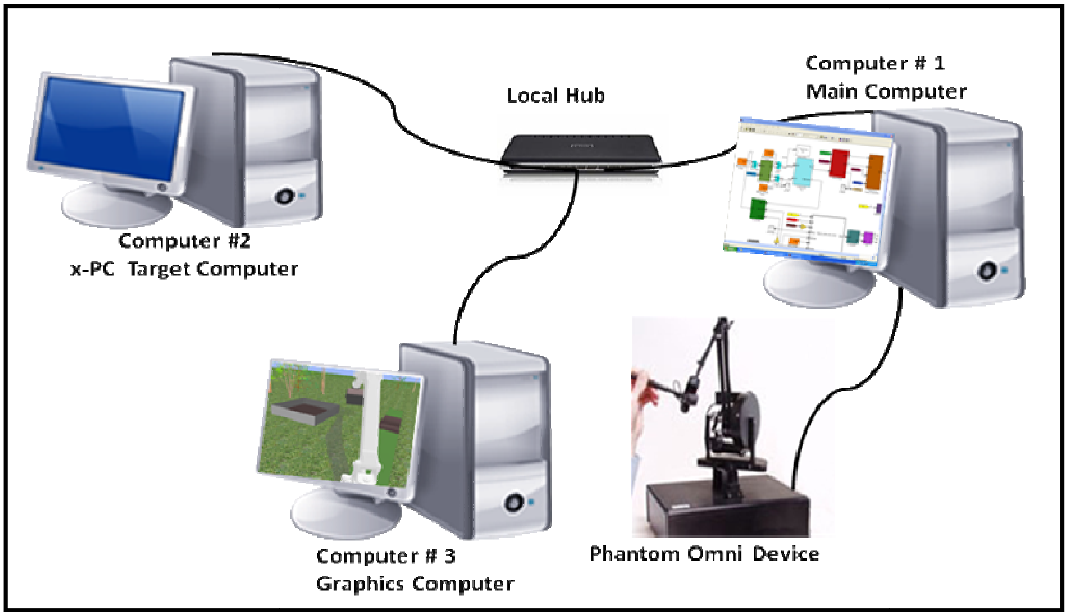

Figure 4.3: Schematic representation of equipment setup

Participants were also briefed on how to complete NASA TLX workload assessment questionnaire, and they were asked to complete one workload assessment after completing each section. Questions about the simulator and controls, and how to complete the workload assessment were answered by the experimenter. Once participants had become familiar with the simulator and its controls and all their questions had been answered, actual testing commenced.

Participants were seated in front of excavator simulation graphics computer (computer #3). They were instructed to use the Phantom Omni device that sat next to computer #3 to control and manipulate the haptic control excavator to fill up bin #1 (left bin) in the simulated work environment. Figure 4.4 shows a participant taking the test, while Figure 4.5 provides a screen shot of the simulated work environment showing the excavator boom/bucket assembly, trench area and the bin.

Each participant had to fill bin #1 under four different conditions (no force feedback, low force feedback, medium force feedback, and high force feedback). These range of force feedback values were determined from the pilot study described in Section 4.5. The Tobii® Eye Tracker was used to record the screen while participants were performing the tasks. This enabled the experimenter to playback each participant's recorded task in order to analyze and extract required data.

Once participants completed the experiment, they were thanked, debriefed and asked to complete a post-test questionnaire. They were also asked for their comments about their experience using the haptic control excavator interface. Overall, the experiment took about one hour to complete.

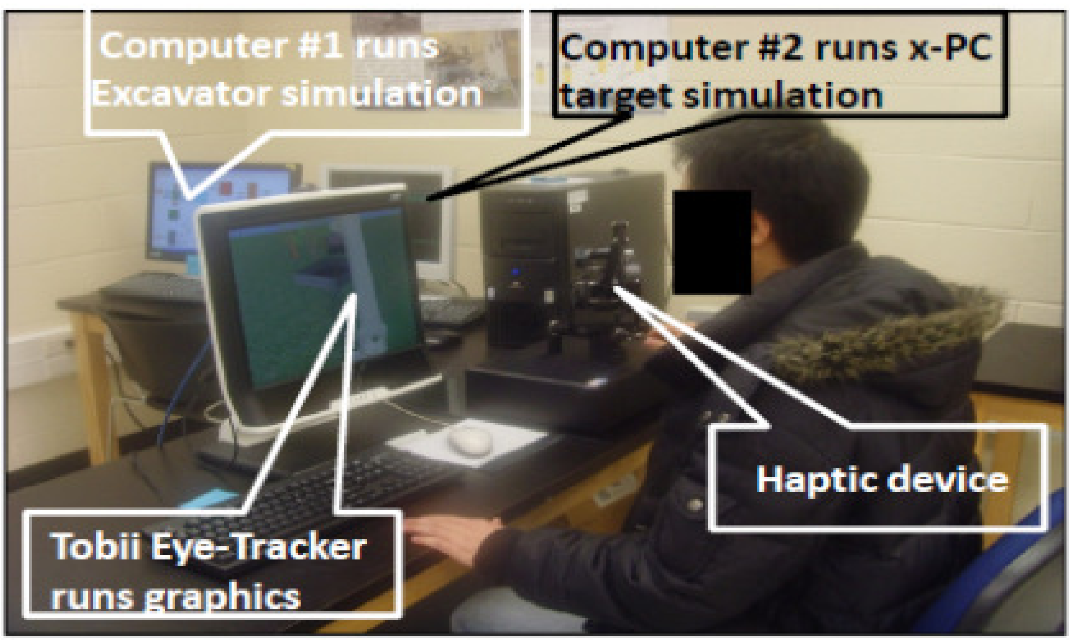

Figure 4.4: Participant taking the test

Figure 4.5: Screen shot of haptic control excavator interface

4.7.6 Data Collection.

The task completion time, number of scoops to fill a bin, and the number of drops per bin were recorded for all participants. Also, a computer based NASA TLX workload assessment questionnaire and a subjective questionnaire were used to gauge participants' subjective assessment of workload under the different force feedback conditions. Appendix B shows task completion time, number of scoops per bin, and the number of drops per bin for each participant.

4.8 Results and Discussions

To compare the performance of operators under different force feedback conditions, task completion time, number of scoops to fill a bin and the number of drops per bin for all participants were compared. Tables 4.5a-c show the mean task completion time, mean number of scoops per bin, and mean number of drops per bin for each of the four force feedback conditions investigated.

Table 4.5a: Descriptive statistics for mean task completion time for each force feedback condition

	No Feedback	Low Feedback	Med Feedback	High Feedback
Mean Task Completion Time (s)	213.96	170.53	150.53	159.59
Std. dev (s)	81.75	64.67	52.92	63.99

Table 4.5b: Descriptive statistics for mean number of scoops/bin for each force feedback condition

	No Feedback	Low Feedback	Med Feedback	High Feedback
Mean # of Scoops/bin	7.45	6.3	6.2	6.3
Std. Dev	1.73	1.08	1.20	1.38

Table 4.5c: Descriptive statistics for mean number of drops/bin for each force feedback condition

	No Feedback	Low Feedback	Med Feedback	High Feedback
Mean # of Drops/bin	0.25	0.2	0.1	0.15
Std. Dev	0.44	0.52	0.31	0.37

Task completion time measured how long it took participants to completely fill up a bin, number of scoops measured the number of times a participant scooped and dumped into the bin in order to fill the it up, and the number of drops measured how many times a participant dropped the content of the bucket outside of the bin. Using appropriate statistical techniques in SAS, the data obtained from the experiment was analyzed to help answer the research questions.

Prior to using any statistical technique, a normality check, test for independence and test for homogeneity of variance (HOV) were performed to ensure no violation. Normality testing using Shapiro Wilk's test revealed violation of normality *(w=0.9175,* and *p=0.0001)* for task completion time, *(w=0.828,* and *p=0.001)* for number of scoops required to fill a bin, and *(w=0.459,* and *p=0.0001)* for number of drops outside of bin. Further, Levene's test for homogeneity of variance showed $(F_{(3, 76)} = 0.55$, and $p=0.6513)$; $(F_{(3, 76)} = 2.38$, and $p=0.0760)$; and $(F_{(3, 76)} = 1.87$, and $p=0.1412)$ for task completion time, number of scoops required to fill up a bin, and number of drops outside of bins respectively. Since the data failed model adequacy test, a non-parametric one-way ANOVA, the Kruskall-Wallis test was used in the analysis.

4.8.1 Research Question 4. Impact of Force Feedback on Performance.

This analysis investigated whether different range of force feedback values affected the performance of operators when using the haptic-controlled excavator interface. The results from the Kruskall-Wallis test showed a statistically significant difference in mean task completion time between the different levels of force feedback *(H=9.94207, 3 d.f,* and *p=0.0242).*

Similarly, as seen from Tables 4.5a-c above, operator performance was affected by the level of force feedback. For example, in terms of task completion time, no force feedback, low force feedback, medium force feedback and high force feedback recorded mean task completion time of 213.96 seconds, 170.53 seconds, 150.53 seconds and 159.59 seconds respectively with corresponding standard deviations of 81.74, 64.67, 52.92 and 63.99 seconds respectively.

Mean task completion time of 213.96 seconds under no force feedback condition improved by about 20.30% under low force feedback condition to 170.53seconds. Under medium force feedback condition, the improvement in mean task completion time was 29.65% (150.53 seconds vs. 213.96 seconds), while under high force feedback condition; the improvement in mean task completion time was 25.41% (159.59 seconds vs. 213.96 seconds) as shown in Table 4.6. Further, there was an improvement of 11.72% in mean task completion time (170.53 vs. 150.53) seconds from low force feedback condition to medium force feedback condition, however, this improvement in mean task completion time diminished when high force feedback was used.

Similarly, the mean number of scoops required to fill up a bin improved as force feedback increased from no force feedback to low force feedback (15.43%) and medium force feedback (16.78%). Figures 4.6-4.8 show mean task completion time, mean number of scoops required to fill a bin, and mean number of drops per bin along with their standard deviations.

Table 4.6: Percentage improvements in operator performance measures

	No Feedback	**Low Feedback**	**Med Feedback**	**High Feedback**
Mean Task Completion Time (s)	213.96	170.53	150.53	159.59
% Improvement	-	20.30%	29.65%	25.41%
Mean # of Scoops/bin	7.45	6.3	6.2	6.3
% Improvement	-	15.33%	16.78%	15.33%
Mean # of Drops/bin	0.25	0.2	0.1	0.15
Mean Error Rate	5%	5%	3%	3%

Further, a one-way non-parametric ANOVA (Kruskall-Wallis) test showed that, in terms of the mean number of scoops required to fill up a bin, there was a significant difference between the different force feedback conditions *(H=9.031, 3 d.f,* and *p=0.0288).* However, a one-way non-parametric ANOVA (Kruskall-Wallis) test showed that, in terms of the mean number of drops per bin, there was no statistically significant difference between the different force feedback conditions *(H=1.6549, 3 d.f,* and *p=0.6470).*

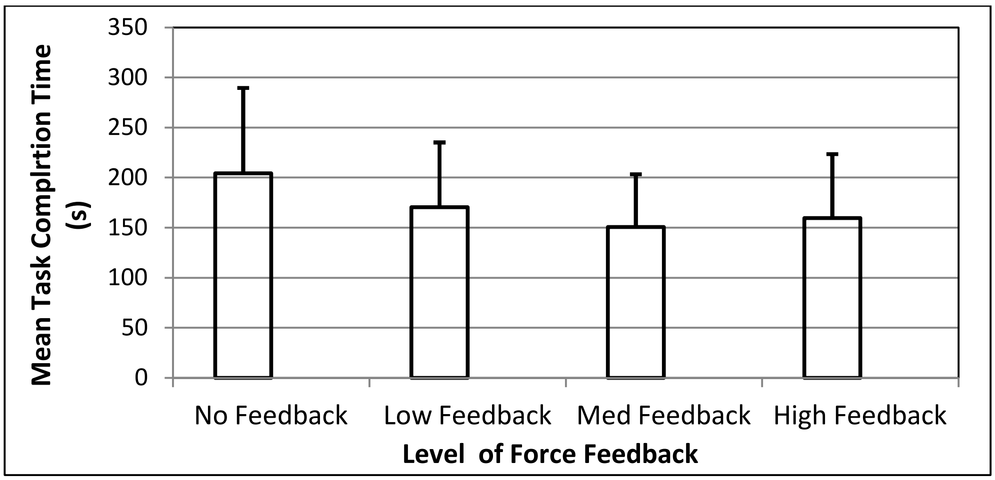

Figure 4.6: Mean task completion time for each force feedback condition

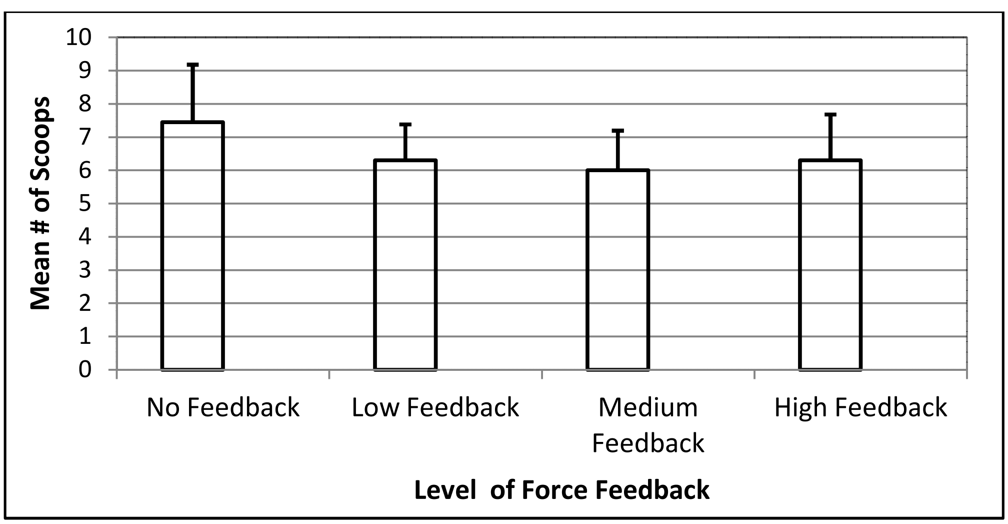

Figure 4.7: Mean number of scoops for each force feedback condition

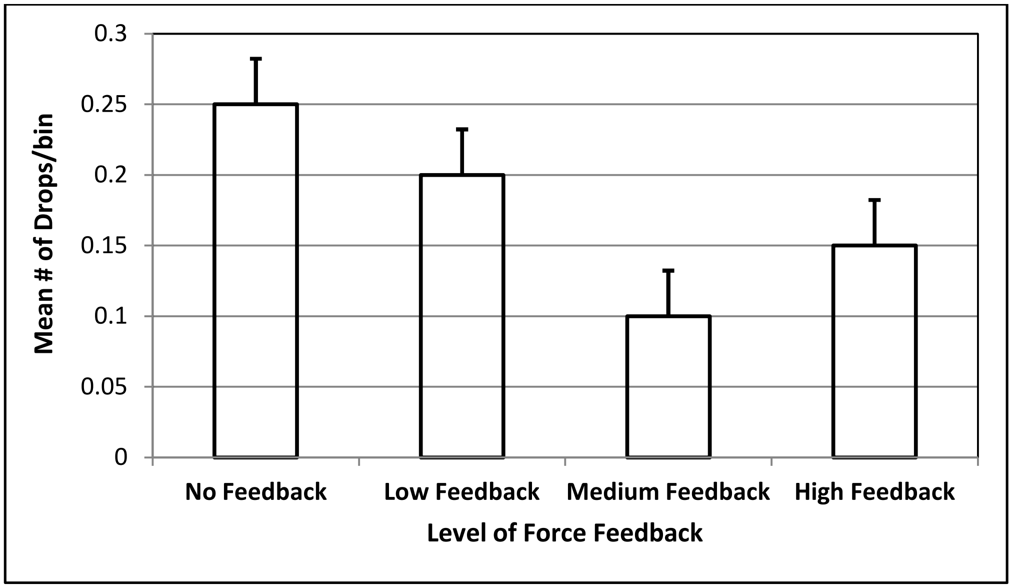

Figure 4.8: Mean Number of drops/bin for each force feedback condition

To compare the means of the different force feedback conditions, a Mann-Whitney-Wilcoxon test was used to compare each pair of the four levels of force feedback for task completion time and number of scoops to fill up a bin. The results of the Mann-Whitney-Wilcoxon comparison test are shown in Table 4.7 and 4.8. The results show that there is a statistically significant difference in task completion time between no force feedback and medium force force feedback *($z=511.00$, $p=0.0097$),* and no force feedback and high force feedback *($z=506.00$, $p=0.0137$),* however, there is no statistically significantly difference between no force feedback and low force feedback *($z=473.50$, $p=0.0963$).*

Lastly, the operator performance under force feedback was compared to performance under no force feedback. The results are summarized in Table 4.9. The results from the non-parametric Mann-Whitney-Wilcoxon test showed a statistically significant difference in task completion time between the no force feedback and force feedback conditions *(z=1012,* and *p=0.0248).*

Table 4.7: Mann-Whitney-Wilcoxon comparison test for mean task completion time

	No Force Feedback	**Low Force Feedback**	**Medium Force Feedback**	**High Force Feedback**
No Force Feedback		z=473.50 *(p=0.096)*	z=511.00 *(p=0.0097)*	z=506.00 *(p =0.0137)*
Low Force Feedback			z=443.50 *(p=0.3775)*	z=438.00 *(p=0.4614)*
Medium Force Feedback				z=400.00 *(p=0.7985)*
High Force Feedback				

Table 4.8: Mann-Whitney-Wilcoxon comparison test for mean number of scoops

	No Force Feedback	**Low Force Feedback**	**Medium Force Feedback**	**High Force Feedback**
No Force Feedback		z=491.00 *(p=0.0300)*	z=500.50 *(p=0.015)*	z=494.21 *(p=0.0251)*
Low Force Feedback			z=425.50 *(p=0.664)*	z=419.00 *(p=0.8060)*
Medium Force Feedback				z=403.00 *(p=0.8516)*
High Force Feedback				

In terms of the number of scoops required to fill a bin, the results showed a statistically significant difference between no force feedback and force feedback conditions *(z=1065,* and *p=0.0029),* however, there was no significant difference in number of drops between force feedback and no force feedback conditions *(z=877,* and *p=0.2446).* The results show that, there was a 25.12% improvement in task completion time under force feedback compared to task completion time under no force feedback. Similarly, there was a 15.84% improvement in the number of scoops required to fill a bin when task was performed under force feedback compared to performance under no force feedback condition.

Table 4.9: Performance under force feedback and no force feedback conditions

	No Force Feedback	**With Force Feedback**	**% Improvement**
Mean Task Completion Time	213.96s	160.21s	25.12%
Mean # of Scoops	7.45	6.27	15.84%
Mean # of Drops	0.25	0.15	25%
Mean Error Rate	5%	3%	2%

In summary, it can be inferred from the analysis above that, the levels of force feedback had a significant effect on operator performance when using the haptic-controlled excavator interface. The results showed a significant improvement of 20.30%, 29.65% and 25.41% respectively for low force feedback, medium force feedback and high force feedback conditions compared to the no force feedback condition.

4.8.2 Research Question 5. Levels of Force Feedback and Performance.

This analysis investigates the range of force feedback values that produce the best operator performance. A summary of the performance measures for the three force feedback conditions are shown in Table 4.10. Mean task completion times were 170.53s, 150.53s and 159.59s respectively for low force feedback, medium force feedback and high force feedback, while the mean number of scoops required to fill up bins were 6.3, 6.2 and 6.2 scoops respectively for low force feedback, medium force feedback and high force feedback.

Finally, the mean error rates were 5%, 3% and 3% respectively for low force feedback, medium force feedback and high force feedback. A one-way non-parametric ANOVA (Kruskall-Wallis) test showed that, there was no statistically significant

Table 4.10: Summary of performance measures for the 3 force feedback conditions

	Low Feedback	**Med Feedback**	**High Feedback**
Mean Task Completion Time (s)	170.53	150.53	159.59
Mean # of Scoops/bin	6.3	6.2	6.3
Mean # of Drops/bin	0.2	0.1	0.15
Mean Error Rate	5%	3%	3%

difference *(H=0.9834, 2 d.f,* and *p=0.6116)* between the means of low force feedback, medium force feedback and high force feedback in terms of task completion time. Similarly, there was no statistically significant difference in mean number of scoops *(H=0.9006, 2 d.f,* and *p=0.2093),* and mean number of drops *(H=0.8539, 2 d.f,* and *p=0.3158)* between low force feedback, medium force feedback and high force feedback.

Further, subjective questionnaire was used to solicit participants' perception of the different force feedback levels investigated. Participants were asked to rate each force feedback levels on frustration, difficulty, comfort and fatigue. The following box plots describe the survey results. When asked to rate which of the force feedback level that was most frustrating to manipulate and control, most participants identified high force feedback as the most frustrating to control. Low force feedback was rated as the least frustrating followed by medium force feedback and no force feedback levels respectively, shown in Figure 4.9. In terms of difficulty and ease of use, participants rated medium force feedback as the most easy to control, while high force feedback was rated as the most difficult as shown in the box plot in Figure 4.10. Finally, when participants were asked to rate force feedback levels in terms of comfort, medium force feedback was rated as the most comfortable, followed by low force feedback. High force feedback was rated the least comfortable followed by low force feedback (see Figure 4.11). Finally, a subjective workload assessment using NASA TLX was conducted to rate participants' perception of workload associated with each of the different levels of force feedback investigated.

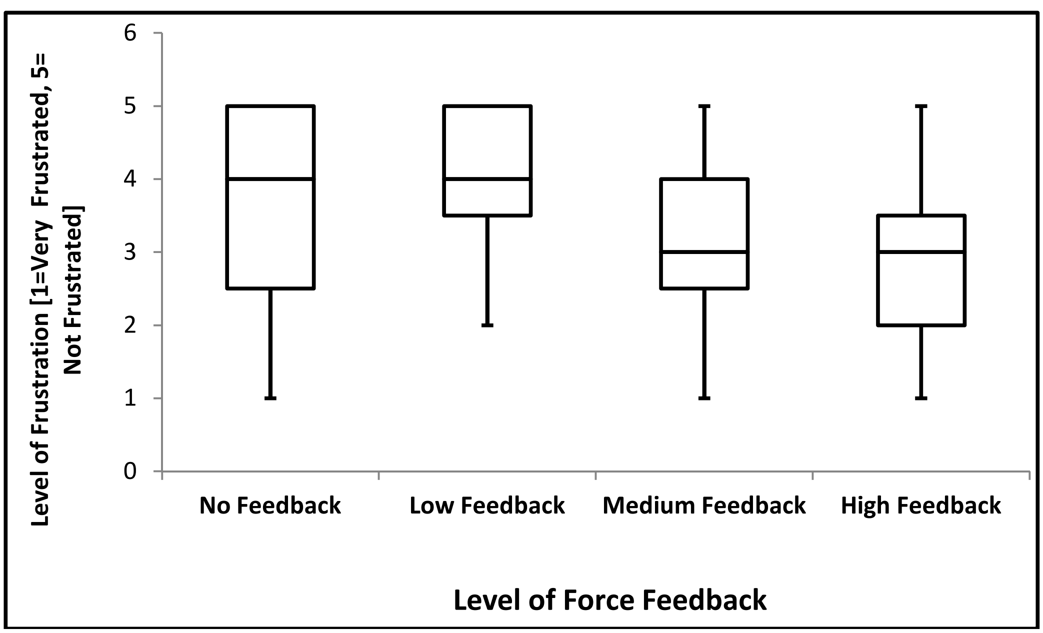

Figure 4.9: Survey results describing participants' level of frustration

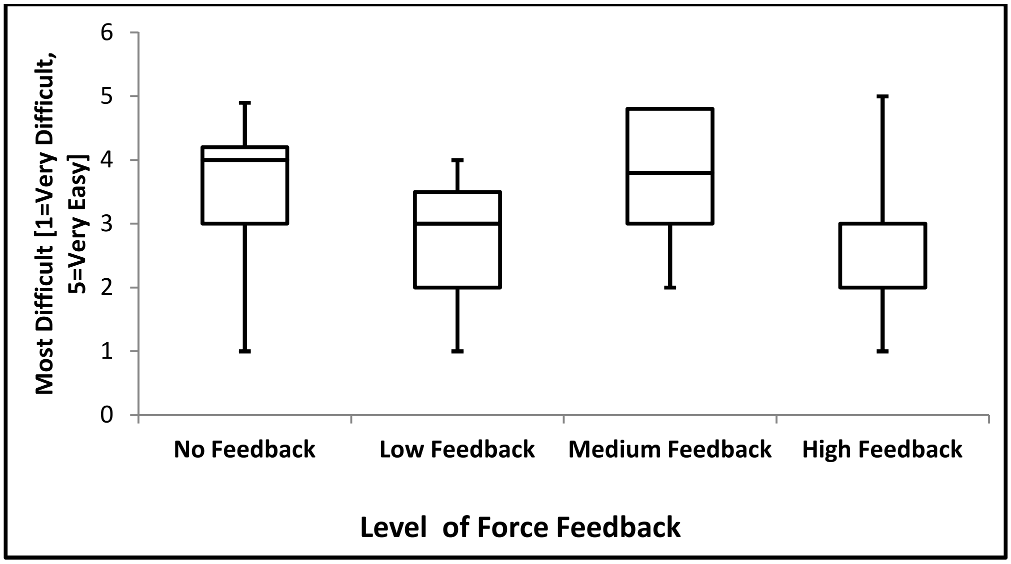

Figure 4.10: Survey results describing level of task difficulty

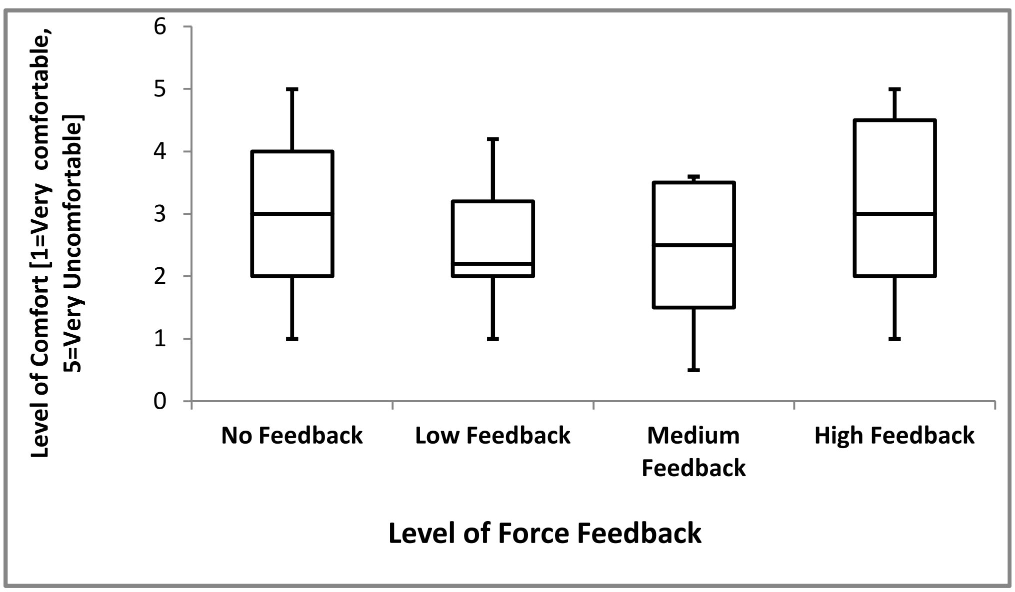

Figure 4.11: Survey results describing participants' level of comfort

The NASA TLX assessment was administered to participants immediately after completing each section of the four excavation tasks. Statistical analysis was performed on the NASA TLX results using levels of force feedback (no force feedback, low force feedback, medium force feedback, and high force feedback) as independent variables, and mental workload, physical workload, temporal workload, performance, effort and total workload as the dependent variables. A plot of the means of the subscales of the NASA TLX ratings as well as the overall workload for each force feedback type is shown in Figure 4.12. The workload scores for no force feedback, low force feedback, medium force feedback, and high force feedback were 65.21 (standard deviation =42.39), 61.32 (standard deviation =22.54), 59.75 (standard deviation =19.49), and 70.92 (standard deviation =15.19) respectively.

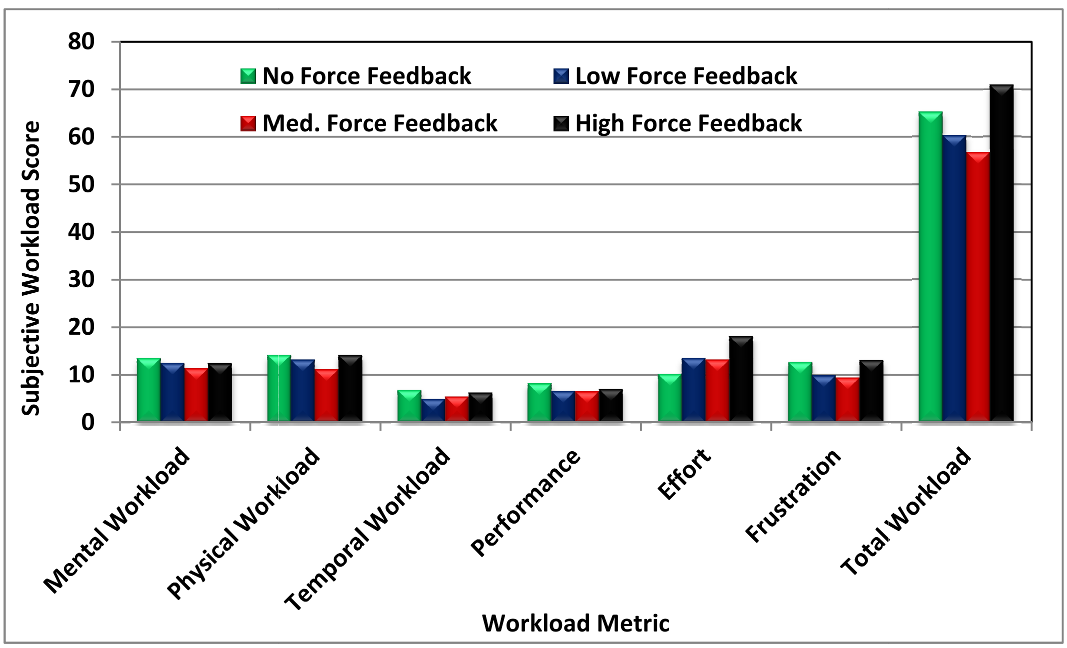

Figure 4.12: Subjective workload ratings for each force feedback type

A normality check of the total workload ratings showed no normality violations (*w=1.047,* and *p=0.083);* therefore, a one-way analysis of variance (ANOVA) was used to analyze the subjective ratings of total workload for the four force feedback conditions. The results from ANOVA showed that, there was no statistically significant difference between the total workload ratings for the four force feedback types ($F_{(3, 80)} = 1.32$, and p

= 0.2743). In terms of the individual subscale measures, physical workload was rated the highest, followed by effort and frustration, while temporal workload was rated the least followed by performance across all force feedback types. This finding was not surprising as participants were under no time pressure to complete the task. The high ratings for physical workload, effort and frustration reflects the fact that participants had to put in physical effort in order control and manipulate the excavator with the haptic device and the frustration they felt due to the poor responsiveness of the haptic device.

Further, comparison of the subscale ratings showed that, medium force feedback had the highest total workload ratings on physical workload, mental workload and frustration. The author believes this may be attributed to the fact that the haptic device was more stable in the medium force feedback range than when other force feedback levels are used. This steadiness allowed participants to control and manipulate the haptic device with the least amount of effort, physical workload and mental workload.

In summary, the results showed that, performance under force feedback was statistically significant compared to performance under no force feedback in terms of task completion time *(z=1012,* and *p=0.0248),* and in terms of the number of scoops required to fill a bin *(z=1065,* and *p=0.0029).* Though medium force feedback range produced higher operator performance in terms of task completion time (150.53 versus 170.53 and 159.59) seconds, as well as the number of scoops required to fill a bin (6.2 versus 6.3 and 6.3), these differences were not statistically significant. However, results from the subjective questionnaire together with NASA TLX results showed that operators rated the medium force feedback range higher in terms of comfort, ease of use, and level of frustration compared to the other levels of force feedback. NASA TLX results also show that the medium force feedback range received the best ratings in mental workload, physical workload and frustration compared to the other force feedback levels.

4.9 Chapter Summary

This chapter presented an empirical investigation to identify the level of force feedback appropriate for use in a haptic control excavator interface. The goal was first to conduct a pilot study to help classify force feedback range values and then conduct experiment to identify which of these force feedback levels produced the best operator performance. Based on the pilot study, four levels/ranges of force feedback were identified based on operator task completion times. These were:

(i) No Force Feedback;

(ii) Low Force Feedback;

(iii) Medium Force Feedback

(iv) High Force Feedback.

An empirical experiment using these force feedback range values identified medium force feedback as the force feedback range with best operator performance in terms of task completion time, number of scoops needed to fill up a bin as well as rate of accuracy. Further, the results show that the level of force feedback affects task performance, for example, task completion time under force feedback improves by about 25.12% compared to task performance under no force feedback condition. Similarly, there is a 15.84% improvement in number of scoops needed to complete a bin under force feedback compared to no force feedback condition. In addition, when using force feedback, the results showed that:

(i) Task completion time under medium force feedback was 15% higher than task completion time under low force feedback, and 10% higher than task completion time under high force feedback.

(ii) There was about 5% improvement in mean number of scoops under medium force feedback compared to both low and high medium force feedback conditions.

(iii) Drops rate improved by about 50% under medium force feedback condition compared to low force feedback and high force feedback conditions.

(iv) Both NASA TLX assessment and subjective questionnaire rated medium force feedback condition higher in terms of operator preference to low force feedback and high force feedback conditions.

(v) Medium force feedback provides a steadier control and, therefore, allowed operators to perform the excavation task more efficiently, compared to the low force feedback and high force feedback conditions. The low force feedback condition was not very useful to operators because, the feedback force in the haptic device was not discernible enough to assist operators, while the feedback force in high force feedback condition made the haptic device very unstable and jittery making it difficult to operate.

CHAPTER 5

QUANTITATIVE MODELING OF HUMAN INTERACTION WITH HAPTIC-CONTROLLED EXCAVATOR

Human actions are partly the results of internal information processing, and since this information flow is internal and invincible, special technologies and methodologies are required to allow inferences to be made and to postulate theories about information flow (Kantowitz & Sorkin, 1983). Modeling how human process information requires a conceptualization of the stages or events that represent the activities and events related to the information. In human-machine interaction, the primary responsibility of the human operator is to extract information (visual, auditory, tactile, etc) for action selection and implementation. The basic structure of human multi-sensory information model (Deng & Ntuen, 1998) is shown in Figure 5.1 below.

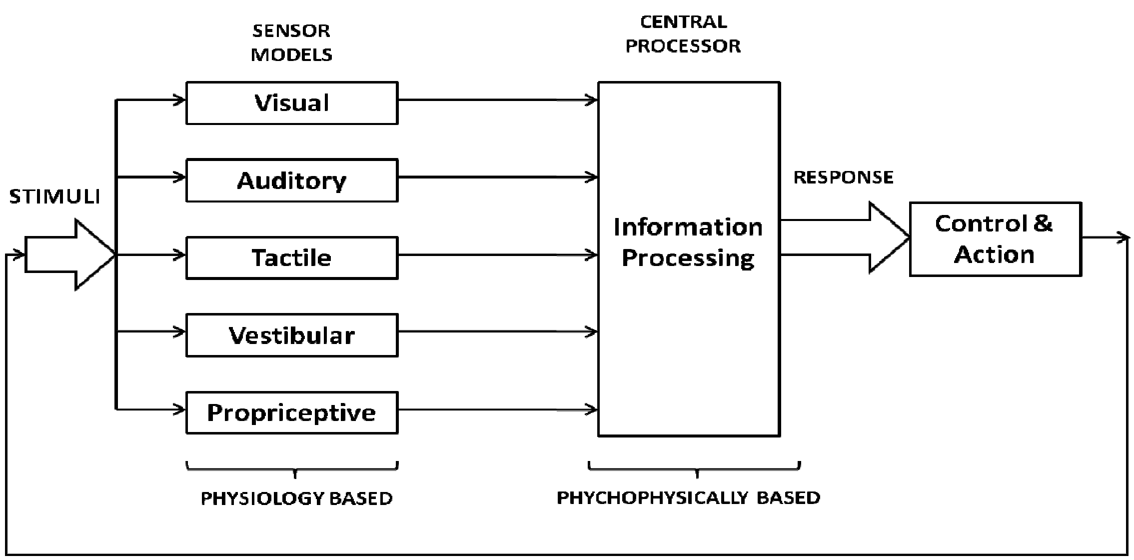

Figure 5.1: Basic multi-sensory model structure [Courtesy of Deng & Ntuen, 1998]

The increased complexity of sensory-based tasks such as excavation requires the human operator to process large volumes of information from multiple sources in order to accomplish the desired tasks. For example, in a typical excavation task, visual, auditory and tactile feedback information may be presented to the operator to enable him/her to accomplish the desired tasks. The feedback information may be presented through the use of a display, touch or auditory technology. Through this, the operator extracts cues which are then sent to the central nervous system for interpretation and necessary action.

In modeling human-machine interaction, control theoretic models have been used because they provide an analytical approach that can describe the actions of humans in a human-machine system (Deng & Ntuen, 1998). Control theory deals with the mathematical analysis of dynamic systems and the mechanisms for achieving a desired state under changing internal and external conditions. Control theory modeling could be classified as either "open-loop" or "closed-loop" depending on whether feedback loop is present or otherwise. A system that has no feedback loop is referred to as open-loop while a system with a feedback loop between the input source and the output node is said to be a closed-loop system. In a closed-loop system as shown in Figure 5.2, output information is fed back to the human operator to help compensate dynamically for errors in the system. In using control theoretic approach to model human performance, the goal is to predict the human performance during task execution.

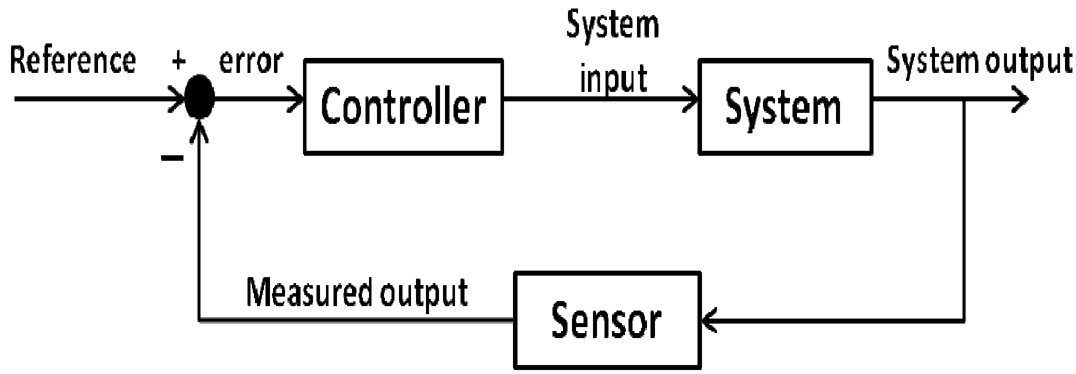

Figure 5.2: Closed-looped control model

5.1 Historical Background to Human Operator Modeling Using Control Theory

Since the 1930s, a lot of research work has been done to study human-in-the-loop man-machine systems. Most of the early work centered on the interaction of the human operator dynamics in aircraft control task for overall aircraft design and control. The desire to analyze aircraft stability, handling qualities and manual control of dynamic systems in a more analytic and mathematical context led early researchers to the concept of control theory, which provided an underlying quantitative theory on which a structured approach to the manual control of aircraft and weapons systems could be developed. One of the pioneers who successfully applied control theory concepts to model human operator dynamic performance was Tustin (1944), who applied control theory concepts to model human control of a power driven gun (George, 2009). He introduced the concepts of quasi-linear systems, describing function, and remnant as applied to the human operator performance modeling of the manual control task. Concepts which compared the control behavior of the human to that of inanimate automatic feedback control system remain key describers today in manual control modeling research and literature, and continue to be applied in human-machine systems design, control analysis, vehicle design, human factor studies, simulator fidelity analysis and vehicle handling qualities research (George, 2009).

Tustin developed describing function of the human operator in a gun tracking task with a delay term $e^{-s\tau}$, a gain component *K*, and a lead term (1+T_Ls) in Laplace transformation as

$$H(s) = Ke^{-s\tau}(1+T_L s) \tag{5.1}$$

where (1+T_Ls) represents operator's tracking performance, $e^{-s\tau}$ represents the overall time delay in operator information processing and response, and K is the operator's principle adjustment parameter.

Since the time of Tustin, several researchers, Elkind (1956), Hess & Shipman (1965), Newell (1967) have spent considerable time and effort to characterize mathematically the dynamics of the human operators which have generated great amount of test and experimental data. These researchers looked at almost every possible control theory approach to model the human operator to different types of vehicles. Other than the aircraft, the automobile is the second most popular vehicle for the application of operator modeling that use control theory, as many of the fundamental concepts and models for the driver could be extended from pilot models.

The most commonly used models that are used to characterize operator performance in manual tracking tasks are the Crossover Model (McRuer & Krendel, 1959), Structural Isomorphic Model (McRuer & Krendel, 1959), Structural Pilot Model (Hess, 1965, 1985), Hosman's Descriptive Pilot Model (Hosman & Stassen, 1999) and the Optimal Control Model (Kleinman et al.,1970). George (2009) in his dissertation provides an extensive review of these models, a brief description of these models is provided below.

5.1.1 The Crossover Model.

Although the crossover model represents the simplest model of manual control task, in modeling human perception, cognition and motor pathways, it provides a surprisingly accurate result for simple manual control problems (George, 2009). The basic Crossover Model from Sheridan and Ferrell (1974) is shown in Figure 5.3.

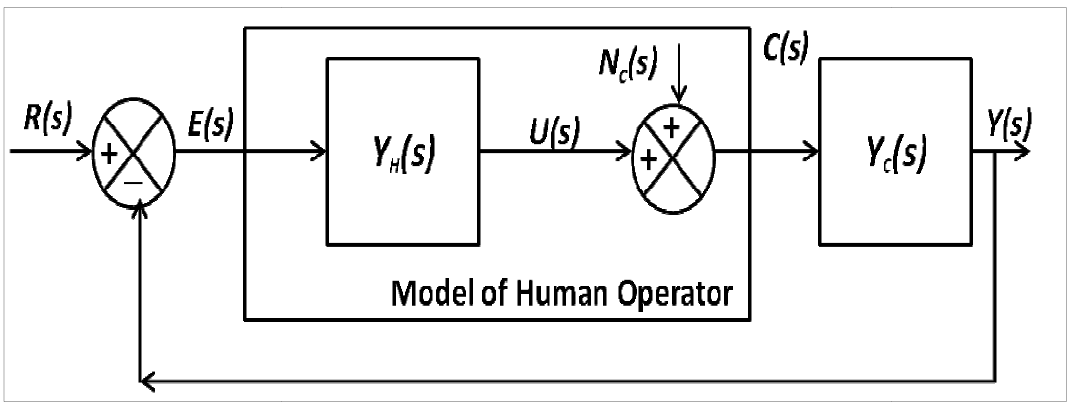

Figure 5.3: Basic compensatory closed loop human operator system model

In its simplest form, the model consists of a human operator describing function $Y_H(s)$ and a linear operator response function $U(s)$ which is diluted by remnant input $N_c(s)$ to produce a total human operator of $C(s)$, which then acts on the machine/plant with machine /plant dynamics $Y_c(s)$. The plant output, $Y(s)$ is then fed back into the reference input and the resulting system error $E(s)$ is received by the human operator describing function to adjust and minimize system error. Using experimental data and simulation analysis, McRuer & Krendel, (1959) developed a relationship for the human operator describing function and the plant dynamics as

$$U(s) = Y_H(s) * E(s) + remnant \tag{5.2}$$

$$\text{where } Y_H(s) = \frac{K_H e^{-\tau_D s}(T_L(s)+1)}{(T_N(s)+1)(T_I(s)+1)} \tag{5.3}$$

K_H is the gain, $e^{-\tau_D s}$is the delay term, lead time is T_L, lag time is T_I and T_N is the first order neuromuscular lag. Lead, T_L and lag, T_I are referred to as the operator equalization terms. McRuer & Krendel (1959) determined through their experiments that for a wide range of basic control models, the vehicle dynamics, $Y_C(s)$ were either K, K/s or K/s^2.

Although the Crossover Model was applicable in a wide range of basic control models, it is not applicable for higher order dynamics modeling especially outside crossover area as it only applies to single degree tracking tasks.

5.1.2 Structural Isomorphic Model.

The full Isomorphic Structural Model is a multiple-path-multiple-feedback model that expands on the crossover model to include all interactions of major perceptual, cognitive and motor pathways, while the reduced Structural Isomorphic Model is a simplified version of the full isomorphic model with many practical applications. It models the human operator using control theory transfer functions of human subsystem behaviors with interpretation of the human psycho-physiological outputs in control engineering terms, and corresponds more to the general model of human behavior with perceptual/sensory, central processing and neuromuscular responses (George, 2009). The Structural Isomorphic Model integrates multiple feedback loops for separate human subsystems to model the overall human operator. The human subsystems include the visual, vestibular, kinesthetic, central processing, proprioceptive, muscle manipulator, spindle/ tendon and nystagmus cross feed. To model total operator dynamics, the

Structural Isomorphic Model attempts to simulate the sensory, central and neuromuscular actuation systems as well as the interaction between the various subsystems, with the assumption that the subsystems add up to the total human dynamic model. This not only allows for total operator investigations, but also a study of psychophysical interactions between subsystems (George, 2009). The original model developed by McRuer (1980) provided a general model that could be adjusted to specific applications depending on the needs of the application such that specific subsystems/channels required for a particular task could be used while those not needed are ignored. McRuer's Structural Isomorphic Model is shown in Figure 5.4 below.

To model the visual system, the structural model uses fovial and peripheral vision as well as eye movement functionality and pathways that carry a continuous representation of display element to enrich field of view. The vestibular and kinesthetic system consists of two sensor types, the semi-circular canals and the otoliths, which are sensitive for angular and linear acceleration, respectively, and are used to model moving human-machine systems such as aircraft and automobiles. The neuromuscular system models proprioceptive feedback/stimulation that the human operator receives from the machine. The central processing system integrates and fuses the visual, vestibular, proprioceptive, and motor functions through cognitive processes and has proven difficult to model due to the complexities associated with modeling the various human sub-systems. This is primarily due to the difficulties associated with determining the parameters of the Structural Isomorphic Model.

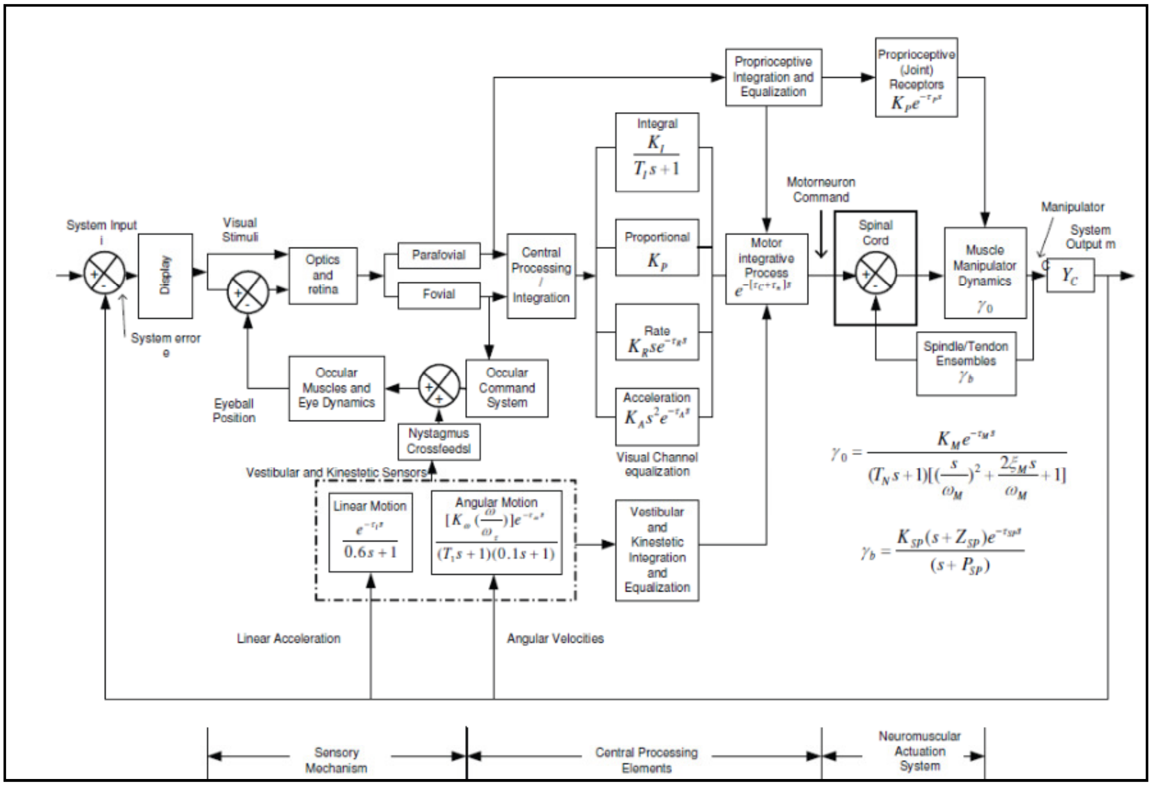

Figure 5.4: McRuer's structural isomorphic model [Courtesy of George, 2009]

5.1.3 Structural Pilot Model.

Hess's original pilot model was derived from a theory put forward by Smith (1976) which proposed that for a closed loop tracking task, the rate control is of fundamental importance to the human pilot, and that rate control is not only important for human-machine performance, but also for operator's perceived vehicle handling qualities. A key point in Smith's theory was that, any model of the human pilot dynamics that structurally corresponds to the human physiology in the tracking task will result in a sound natural and physical measure of pilot handling quality assessment (George, 2009). Smith's human operator model is shown in Figure 5.5, where Y_H represents the structural pilot model, Y_C represents the control elements; *m* is the control output, *C* is the reference input, and e is the pilot perception error. Given that the human transfer function Y_H is given by $Y_H = \frac{K_e}{K_m Y_c s+1}$, then the open loop transfer function for simple gain K, integrator (velocity control) K/s, and second order (acceleration) control K/s^2 are

$$Yc = K,\ Y_H Y_C = \frac{K_e K}{K_m K s+1} \tag{5.4}$$

$$Y_C = K/s,\ Y_H Y_C = \frac{K_e K}{s(K_m K_s+1)} \tag{5.5}$$

$$Y_C = K/s^2,\ Y_H Y_C = \frac{K_e K}{s^2\left(\left(\frac{K_m K}{s}\right)+1\right)} \tag{5.6}$$

By using time delay components, and adjusting the gains, it is possible to derive the crossover model from Smith's model.

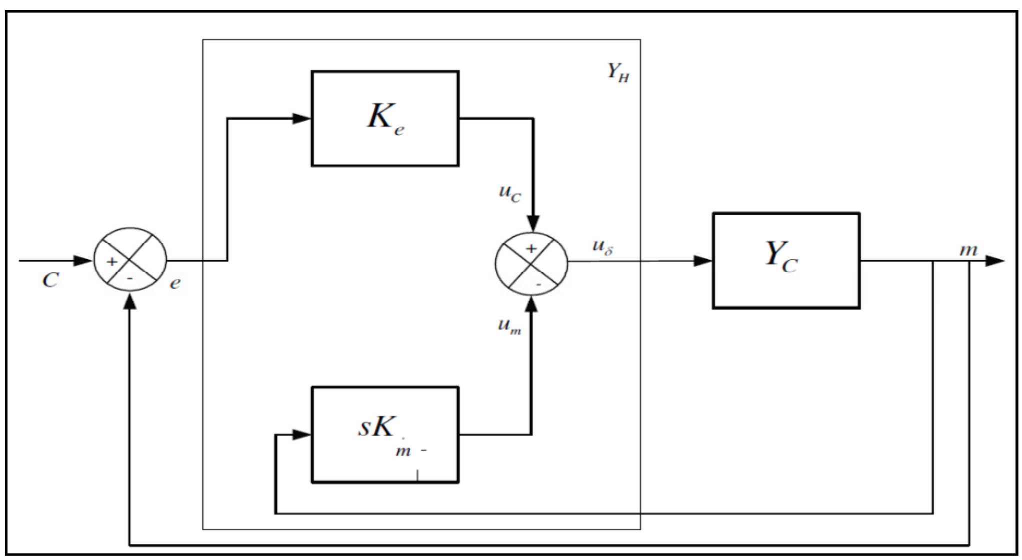

Figure 5.5: Smith's human operator model [Courtesy of George, 2009]

Hess's structural pilot model is shown in Figure 5.6. In this model, *Ke* represents the gain of error e, *Km* is the effect factor for the output feedback m, *Kc* is the gain for input c, K_1 and K_2 are the gains for the proprioceptive feedback loops, T_1 and T_2 are the time constants for Y_f and Y_m; τ_0 and τ_1 are the time delay constants for the error and input signals; ξ_n is the damping ratio and ω_n is the crossover frequency.

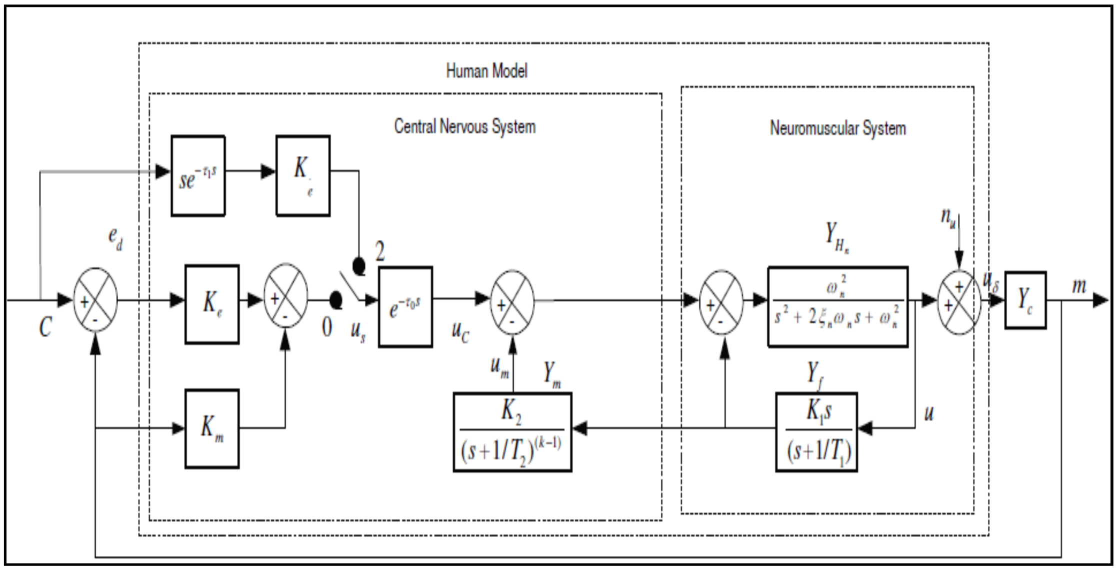

Figure 5.6: Simplified structural pilot model [Courtesy of Rouse, 1995]

5.1.4 Descriptive Pilot Model.

Hosman's descriptive pilot model is shown in Figure 5.7. In the descriptive pilot model, sensors represented by transfer functions, are placed in parallel to convert the stimuli, attitude, angular rate, and angular acceleration, to the sensory outputs $R_i(\omega)$. Sensory information is integrated into a single output in the central nervous system (represented by the summing block in the model) where each of the modalities is weighted by a weighting factor W to define the contribution of that particular sensory input. A detailed and comprehensive review of the descriptive pilot model can be found in (Hosman & Stassen, 1999).

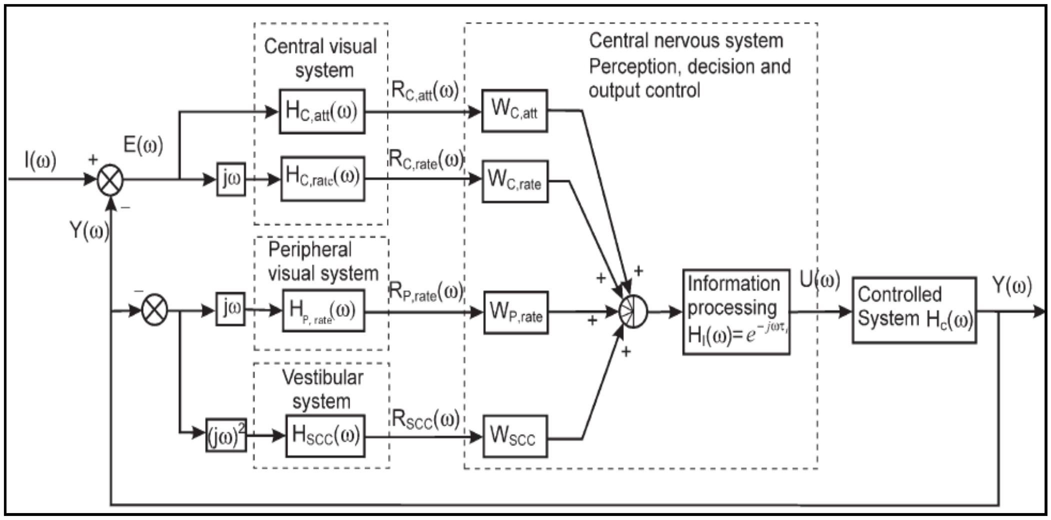

Figure 5.7: Hosman's descriptive pilot model

5.1.5 The Optimal Control Model.

Kleinman & Baron (1970) proposed optimal control for the analysis of human-machine systems, by incorporating the concepts and components of McRuer's work (such as time delay, remnant concept and neuromotor model). To model human-machine system, Kleinman & Baron (1970) employed state space concepts, Linear Quadratic Gaussian optimal control theory and estimation theory to generate human operator models (George, 2009). The basic assumptions invoked in modeling human-machine system using the optimal control modeling are

(1) That the human is well-motivated and behaves in an optimal manner subject to his inherent limitations and to the requirements of the control task
(2) The human has an accurate internal model of the system dynamics and parameters affecting his control behavior
(3) The human has the expertise required for the control task

These human characteristics are subject to operator psychophysical limitations and constraints such as time delays, system remnant and neuromuscular dynamics. Humans can adapt to environmental changes by using natural sensors to receive and process information, and make real-time decisions based on their feedback mechanisms. This human trait has made the application of control theory to human performance modeling attractive to many multimodal researchers (Deng, 1999).

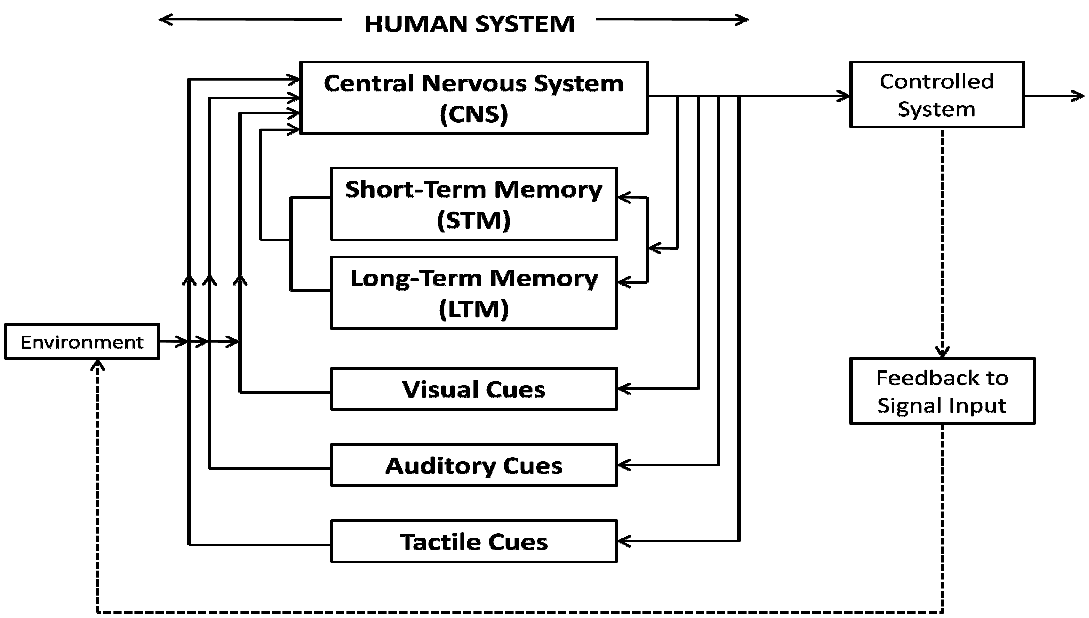

Figure 5.8: Conceptual architecture for multi-sensory information processing
[Courtesy of Deng, 1999]

The commonly observed and cited rationales as outlined by Deng (1999) are as follows

(1) The human behavior changes with respect to time and space. The spatial attributes include such things as external and internal stimuli, the intensity of the stimuli, and the boundary of the human-system interaction.

(2) The human-system interaction is subject to some perturbations such as environmental noise during task execution.

(3) The human operator relies on his/her sensory information for feedback. Because the human behaves as a sub-optimal learner, he/she attempts to select an input-output information channel to improve performance. (Rouse, 1980).

(4) The human is considered to be an intuitive statistician, a property that allows him to filter and smooth relevant information after data has been observed. The filtered data is then used for control decision making.

These properties of the human operator allow his task performance and behavior to be modeled using control theory (Fang, 1997). Further, humans as operators are known to behave in ways that tend to minimize their errors when performing control tasks by using feedback information from prior tasks. The classic human control model for manual control tasks developed by McRuer & Krendrel (1974) for human pilots in pursuit compensatory tracking tasks is shown below.

$$Y_p = \mathrm{K_p}\, e^{-s\tau} \left(\frac{T_{Ls}+1}{T_{Is}+1}\right) \left[\frac{T_{Ks}+1}{T'_{Ks}+1}\right] \left\{\frac{1}{(T_{N1S}+1)\left[\left(\frac{S}{\omega_N}\right)^2 + 2\frac{\xi_N}{\omega_N}S+1\right]}\right\} \qquad (5.7)$$

where K_p is Gain, $e^{-s\tau}$ is pure information transmission time delay, and τ is estimated to range from 0.06 to 0.1 sec., $\left(\frac{T_{Ls}+1}{T_{Is}+1}\right)$ is series equalization,

$\left[\frac{T_{Ks}+1}{T'_{Ks}+1}\right]$ is the low frequency lag-lead, and $\left\{\frac{1}{(T_{N1S}+1)\left[\left(\frac{S}{\omega_N}\right)^2 + 2\frac{\xi_N}{\omega_N}S+1\right]}\right\}$

represents the dynamics of the neuromuscular actuation system of the arm with typical values $1/T_{N1} = 10sec^{-1}$, ω_N = 16.5rad/sec, and ξ_N = damping coefficient.

The principle idea here is that the human operator adapts a behavior characteristic to the environment and behaves like a "good servo" in the region of the crossover frequency resulting in a constant overall open-loop transfer function for the system (McRuer, 1967). The open-loop transfer function is then given by

$$G_0(s \approx jw_c) = G_h(s)*G_p(s) = \frac{K_c e^{-\tau_c s}}{s} \tag{5.8}$$

where $G_h(s)$ is the transfer function modeling the human behavior as a linear feedback controller and $G_p(s)$ is the system transfer function, K_c = speed control, K_c/s = heading control at low to moderate speeds, and $e^{-\tau_c s}$ is pure information transmission time delay.

In humans, the central nervous system (CNS) is the central mechanism through which information processing takes place. The central nervous system performs information filtering and integration tasks based on motor commands it receives from neuromuscular system. Visual, auditory and tactile cues from the environment are encoded in different frames of reference than those in the coordinates of the central nervous system, and then sent to the extra-ocular muscle along a final common pathway (Deng, 1999). By combining the functions of the central nervous system and the neuromuscular system, McRuer (1974) was able to represent the human operator system by a transform function given by

$$G(s) = \frac{K_H e^{-T_e s}(1+T_L s)}{(1+T_N s)(1+T_I s)} \tag{5.9}$$

5.2 General Structure of the Optimal Control Models

The general structure of the optimal control model developed by Kleinman *et al.*, in 1971 is shown in Figure 5.9.

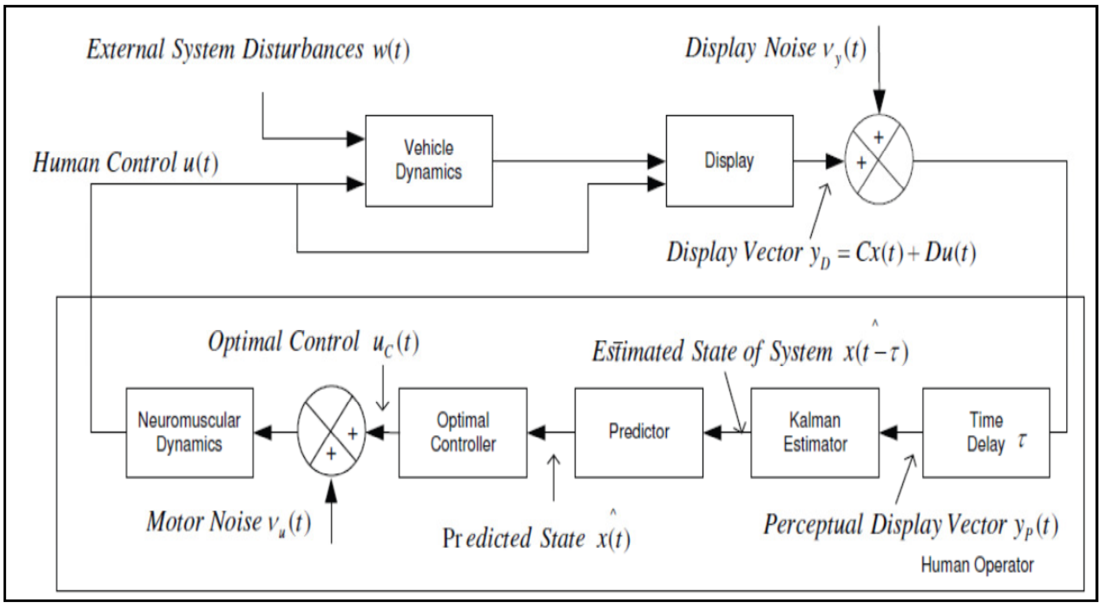

Figure 5.9: Optimal control model of human operator

The parameters of the optimal control model as outlined by Hess, 1976 are listed below.

1. Time delay: A pure time delay is included in each of the control models.
2. Neuromuscular dynamics: Each output of the neuromuscular system is modeled as a first-order lag.
3. Observation and motor noise: Each variable which the human operator observes from his display is assumed to contained a human-induced additive noise related to the variance of the observed cue.

4. Rate perception: If a stimulus is perceived explicitly, then the human perceives the first order derivative of the sensory stimuli but not higher derivatives, and the displayed stimuli is also noise contaminated.
5. Index of performance: the index of performance is subjectively selected by the human to minimize error rate.

The optimal control model is generally represented by the state equation given by

$$\dot{x}(t) = Ax(t) + bu(t) + w(t), \tag{5.10}$$

where, x(t) is an n-dimensional vector representing the random input, $u(t)$ is a scalar representing the human input system, and $w(t)$ is an n-dimensional vector representing random disturbance to the system. A is an nXn matrix and b is an $nX1$ matrix. It is assumed that the system is completely controllable, and that at the minimum, one system output can be described by the equation

$$Y(t) = cx(t) + du(t) \tag{5.11}$$

where $Y(t)$ is system output, c is an mXm matrix and d is an $mX1$. In display applications, when viewing a display foveally or peripherally, the operator perceives an output that is both time-delayed and noise corrupted as observed by (Kleiman *et al*., 1971). Hence, output y_p can be described by the equation

$$y_p(t) = y(t-\tau) + v_y(t-\tau) \tag{5.12}$$

Therefore, system output equation as described in equation *(5.12)* can be written as

$$y_p(t) = cx(t-\tau) + du(t-\tau) + v_y(t-\tau) \tag{5.13}$$

Equation *(5.14)* is referred to as the delayed noise version of equation *(5.12)*. This is the signal that is processed by the operator in order to yield the command input $u_c(t)$.

Further, a motor noise given by $u_m(t)$ representing the random error in control execution results due to operator's poor knowledge of the system input $u(t)$ is added to $u_c(t)$. Hence,

$$u(t) = u_c(t) + u_m(t) \quad (5.14)$$

Further, (Deng, 1999) proposed an optimal control model for multimodal information processing formulated as a multiple input-multiple output (MIMO) model. The general block diagram of Deng's optimal control model is shown in Figure 5.10 below. As in other human operator control models, Deng's model assume that multimodal information processing tasks are governed by the optimal control model equation given in *(5.15)* as

$$\dot{X}(t) = AX(t) + Bu(t) + Dw(t) \quad (5.15)$$

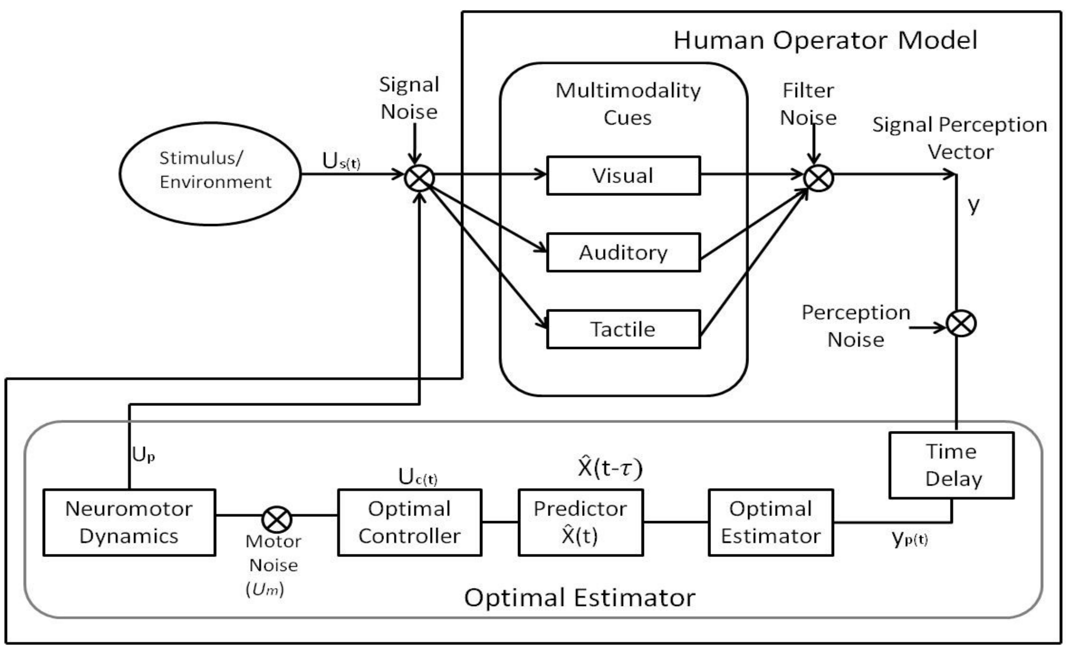

Figure 5.10: Optimal control system for multimodal system

5.3 Conceptual Model of Human-Excavator Interface

In order to model and simulate the processes involved when a human operator interacts and manipulates an excavator in order to accomplish an excavation task, it is necessary to conceptualize the processes that describe the interaction. The operator uses the central processing system (long term memory and short term memory) to process information/cues received from the environment and then decides on the necessary action to take. The operator's response is then executed in the form of commands given to the excavator, e.g. moving or rotating the haptic device in order to accomplish a given required task. A conceptual framework for the interaction between the operator and the excavator is shown in Figure 5.11.

The model has the following interacting components: (i) the human sensory system, (ii) the human output modalities, (iii) excavator input channels, (iv) excavator output modalities, (v) the central processing system, (vi) the excavator processing system as well as the visual, haptic, auditory and force feedback cues.

(i) Human sensory system: The human sensory/input system includes human senses that the operator uses to receive feedback/information from the excavator and the environment such as the eyes, ears and the skin *(palm)* for sensing vibrations. *Force feedback* that operator receives allows a simultaneous exchange of information between operator and excavator and results in a more *immersed* interaction between operator and machine.

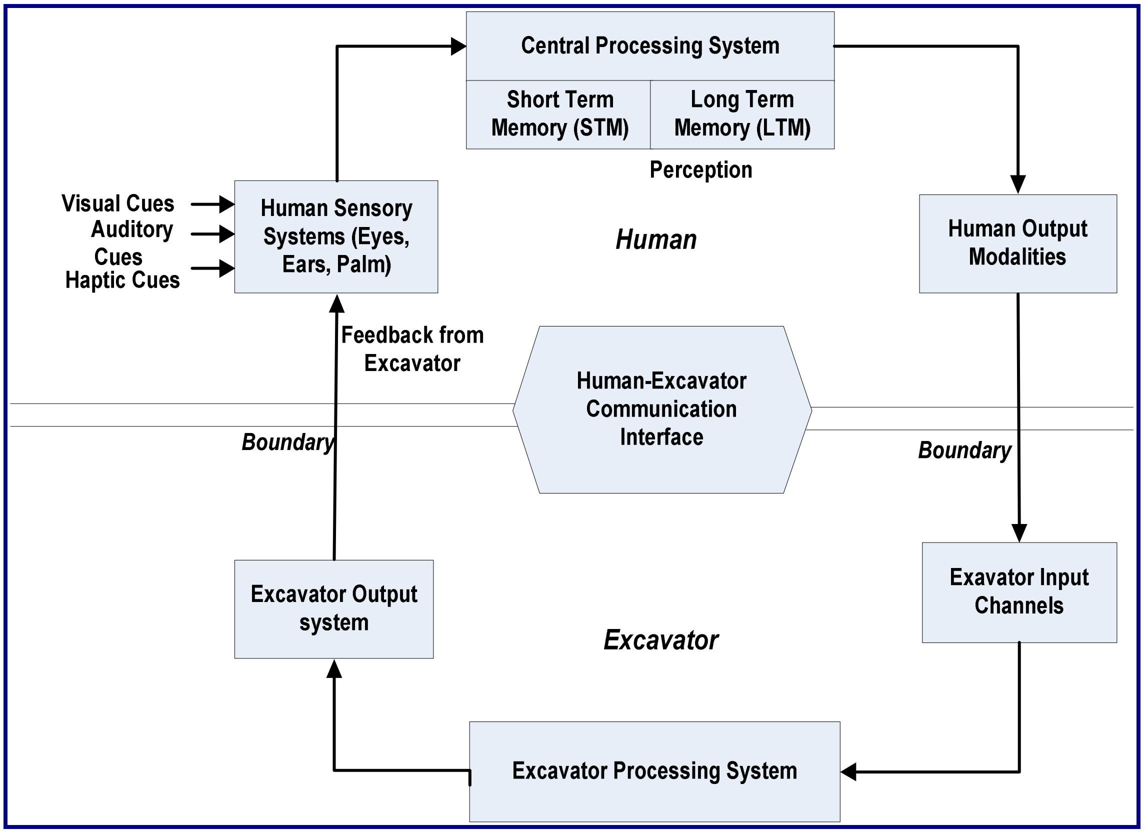

Figure 5.11: Conceptual model of operator-excavator interaction

(ii) **Human output modalities:** The human output modalities include operator actions such as touch, gaze (e.g. stylus rotation) that are used by the operator to control and manipulate the excavator. The output modalities (operator actions) are influenced by environmental cues as well as force feedback from excavator. For example, an operator may decide to stop scooping task if force feedback alerts him to the presence of obstacles/foreign materials. Thus, there is an interaction between the actions that the operator takes and environmental as well as feedback cues.

(iii)**The excavator input channels:** The excavator input channel includes the levers, stylus etc. through which the excavator receives commands from the user. These commands are then processed through the built-in mechanical/electrical architecture to produce mechanical actions such as bucket and boom movements.

(iv) **The excavator output modalities:** The excavator output modalities provide feedback from the excavator to the operator in the form of force feedback (vibration), audio or visual feedback. Common audio feedback that operator receives from excavator include audio alerts/alarms or engine revving sound though it may require some level of experience to fully utilize sound from revving engine.

(v) **Excavator processing system** includes the built-in mechanical and electrical architecture/algorithms that allows excavator to convert operator commands (e.g. stylus rotation) into mechanical actions (e.g. bucket open/close) and be able to generate force feedback to operator.

(vi) **The central processing system** of the operator include resources from both short term memory and long term memory that the operator uses to perceive cues from the environment, process the perceived cues and make decisions by issuing commands to the excavator (through stylus manipulation) in order to accomplish the required task. It must be noted that the interaction between these components takes place simultaneously rather than sequentially, though components are treated separately for the purpose of graphical display.

5.4 Developing a Model for Human-Excavator Interaction

The human-excavator system is a complex system with high degree of freedom, however, for ease of modeling, and to reduce the level of complexity into a manageable level, the subsystems of the human operator are assumed to be linear. To determine a model for the human-excavator interaction, transfer function for each dynamic component of the human subsystem is obtained as measurable input-output relation. A graphic representation of the human-excavator interaction and feedback loop is shown in Figure 5.12.

The goal of this modeling work is primarily concerned with how the excavator operator combines and integrates visual, auditory and haptic cues/signals from the task environment to improve his/her performance on the task. Therefore, even though the dynamics of the haptic device and the excavator along with their complexities are essential to how the operator interacts with the excavator in general, it is not the focus of this model.

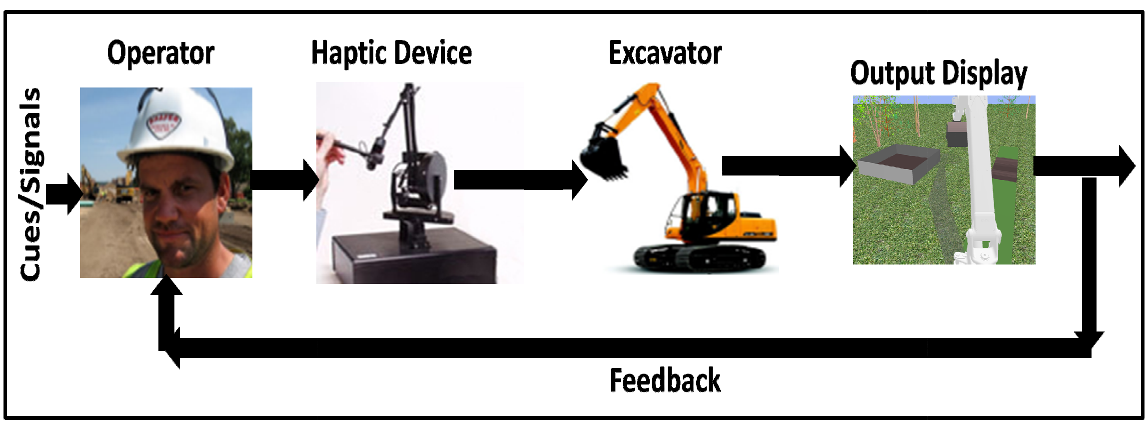

Figure 5.12: Components of haptic-controlled human-excavator model

To develop a quantitative model for the interaction between operator and the excavator, the human-excavator system is assumed to be a closed-loop control system. For ease of modeling and implementation, the model is assumed to be a simple closed-loop system as shown in Figure 5.13. In this model, the human operator receives signals/cues, *R(s)* (made up of visual, auditory and haptic cues) with probabilities ρ_v, ρ_a and ρ_h respectively from the environment. The operator processes this information in the central processing system, sends signals to neuromuscular system which then acts on the haptic manipulator device with input *U(o).* The haptic manipulator device is also acted on by manipulation noise *w(t),* and sends output *U(o)s* to the excavator system dynamics. The excavator performs the excavation task and the output *Y(s)* is corrupted by display noise $V_y(t)$. The displayed output $Y_d(s)$ is fed back to the operator, and it is subtracted from the input signals *R(s),* which results in error term *e(t)* that the operator minimizes in order to improve his/her performance on the task. In this model, the excavator dynamics is represented by the state equations *Y*(s), and inputs include human control *U(o)s* and manipulation noise *w(t).*

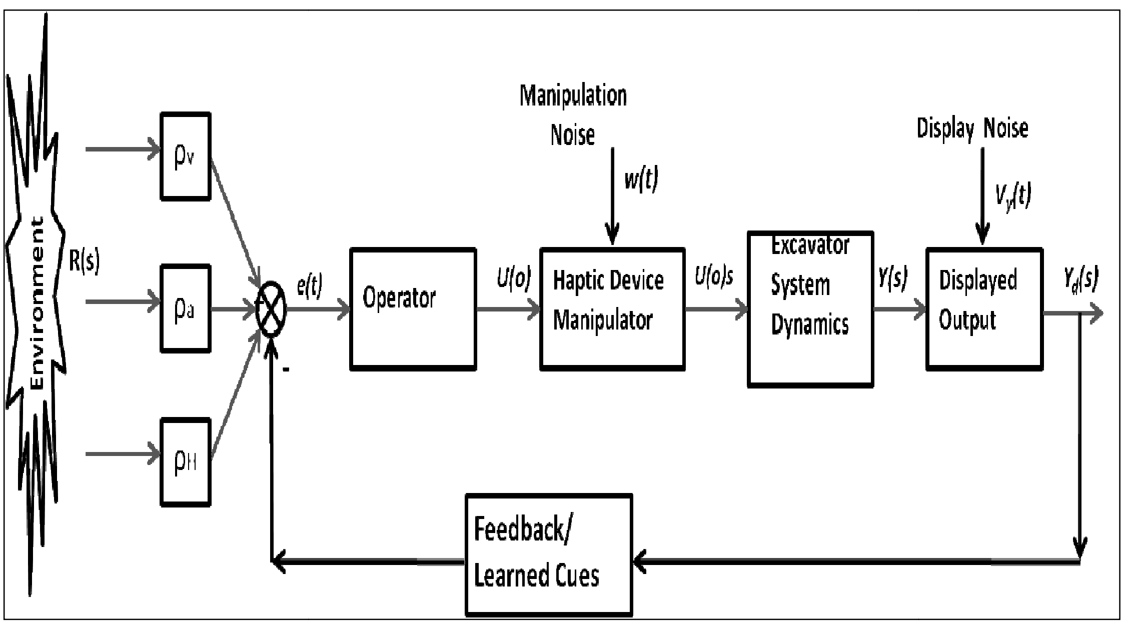

Figure 5.13: Haptic-controlled human-excavator model

Manipulation noise *w(t)* and display noise $V_y(t)$ accounts for operator limitations in perceptual resolution, image processing, situational awareness and attention switching. A structural representation of the human operator in multi-sensory excavation task is shown in Figure 5.14. As shown in the model, signals can come from visual cues, auditory cues or tactile cues. Since the human operator only directly observes the output of the system rather the system states, it is important to include errors to account for the difference between the observed versus the system state. Depending on the direction of deviations between the observed and system state, feedback could be positive or negative. The visual, auditory and haptic cues with known transfer functions Y_v, Y_a and Y_h respectively are integrated based on how a particular cue contributes to overall perception. When the cues arrive at the central nervous system (CNS), they are first processed by the memory filter with a known transfer function.

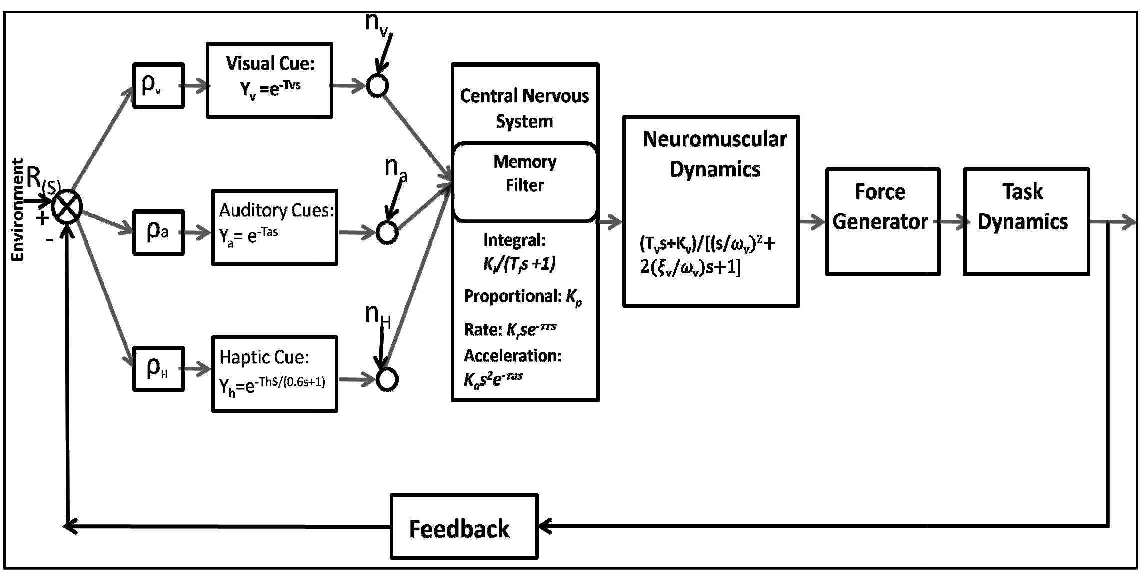

Figure 5.14: Structural model of human operator in multi-sensory excavation task

Five transfer function blocks are identified for the model. These are the signal input transfer function, the central nervous system transfer function, the neuromuscular dynamics transfer function, the force generator transfer function and the task dynamics transfer function. Each of the subsystems in the structural model is described below.

5.4.1 Signal/Cue Input and Perception.

Visual cues are the primary source of information for human operator in most control tasks. This information could be derived from displayed instruments, out-the-window view or a combination of both. An important feature of the visual pathway is its ability to provide a continuous display of signal to the operator by virtue of parallel fovea and parafovea pathways even when the eye is scanning, visual cues can be represented by the transfer function

$$Yv = e^{-\tau_v s} \qquad \textit{(5.16)}$$

with time delay τ_v values between 140ms-300ms (McRuer,1980) . Also, auditory cues are considered as processing time delays in control model similar to visual cues (Hess, 1995) with transfer function

$$Ya = e^{-\tau_a s} \qquad \textit{(5.17)}$$

with time delay τ_a=210ms often used to approximate the auditory information processing time delay. Further, according to McRuer (1980), tactile cues can be perceived as linear acceleration with transfer function

$$Yh = \frac{e^{-\tau_h s}}{0.6s+1} \qquad \textit{(5.18)}$$

with time delay τ_h = 0.1s. Using the structural model of the human operator in multi-sensory information processing task shown in Figure 5.14, a transfer function for sensory cue integration can be developed for the human operator in a task that requires visual, auditory and haptic information (McRuer, 1980; Deng, 1999). Given that Y_v, Y_a and Y_h are the transfer functions of visual, auditory and haptic cues, and ρ_v, ρ_a and ρ_h are the probabilities of occurrence respectively for visual, auditory and haptic cues, the transfer function of sensory cues integration $G_{(CUES)}$ is given by

$$G_{CUES}(S) = \rho_v Y_v + \rho_a Y_a + \rho_h Y_h \qquad (5.19)$$

where ρ_v is probability of visual cues, ρ_a is probability of auditory cues, ρ_h is the probability of haptic cues, and $\rho_v + \rho_a + \rho_h = 1$.

5.4.2 Central Nervous System (CNS).

The central nervous system is the primary source for human information processing, and integration and sensory fusion of visual cues, vestibular, proprioceptive and motor functions all occur here. It performs information integration and fusion based on motor commands and neuromuscular inputs. Visual, auditory and haptic sensory data are all initially encoded in various frames of reference that are different from the coordinate system by the central nervous system. It is then sent to the extra-ocular muscles along a final common pathway, where it is combined with other current information to direct the eye to the target location (Deng, 1999). Several mathematical models Hess (1985) have been developed to quantify the processes that occur in the central nervous system. The central nervous system processes have been modeled as integral effect, proportional effect, rate effect, and acceleration effect. Proportional and rate effects are modeled as time delays while acceleration and lag effects are modeled as latency differences. Equations *5.20a-d* represents the integral, proportional, rate, acceleration effects respectively. For this work, the proportional model of the central nervous system will be used.

$$G(s) = \frac{K_I}{T_I s+1} \tag{5.20a}$$

$$G(s) = K_P \tag{5.20b}$$

$$G(s) = K_R s e^{-\tau_R s} \tag{5.20c}$$

$$G(s) = K_A s^2 e^{-\tau_A s} \tag{5.20d}$$

5.4.3 Neuromuscular Dynamics System.

The central nervous system (brain and spinal cord) and neuromuscular system (nerves and muscles) form the nervous system. Muscular activities that are associated with tracking and manipulation are characterized by changes in the length and tension of the antagonist/agonist muscles pairs that drive these muscles. Muscles are mostly made out of "*muscle fibers*" called myofibrils whose chemical structure allows muscles to contract. To enable bones to move in multiple directions, antagonistic pairs of muscles are often present in the human body. When one of these muscles (the flexor) contracts, tension on the bone increases, and allow rotation in one direction. When the other muscle (the extensor) contracts, tension on the bone is released, enabling motion in the other direction. Changes in tension and rate of tension can be sensed by the Golgi tendon organs while changes in length and rate of change of length are sensed by the muscular spindles.

By assuming that the human-excavator system is operating on random-appearing signals with stationary statistics, the neuromuscular system could be assumed to be fluctuating about an operating point corresponding to some steady-state or average tension. Thus, muscle contractions which allow rotation in only one direction can be assumed as either positive or negative fluctuations of the agonist/antagonist pairs about a steady tension bias value and permits neuromuscular dynamics to be greatly simplified and modeled (McRuer, 1980). As shown in Figure 5.13, highly simplified model of the neuromuscular dynamics of the human operator is given by the transfer function in Equation 5.21.

$$G_N(s) = \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega n}\right)^2 + 2\left(\frac{\xi n}{\omega n}\right)S+1\right]} \tag{5.21}$$

where $K_n = 2$, $T_n = 0.2$, $\omega_n = 20$, and $\xi_n = 0.825$

Operator's neuromuscular response is modeled as a first order lag and represents the physical limitation on the operator's overall working ability, as well as the subjective constraints associated with a good/bad task execution. For example, a good operator will make few erratic and rapid control inputs.

5.4.4 Force Generator.

The force generator accounts for the intentional force that is generated as a result of the operator's reaction to sensory feedback which forms an external feedback loop with the human operator in the loop (Deng, 1999). When performing a tracking or manipulation task, the human operator internally generates models appropriate for the task after receiving visual, auditory or tactile feedback cues. It is, therefore, important to include the operator's internal mechanism necessary to generate the trajectory of operator's reaction to both visual and tactile feedback when developing a model of the human operator (Deng & Ntuen, 1998). However, for the human-excavator model under consideration, the force generator is assumed to be a constant and equal to one (1).

5.4.5 Task Dynamics.

The task dynamics in the model represents the performance of the excavator. Operator control behavior changes drastically when higher order controlled system is encountered. For example, with simple gain K or integrator K/s, the operator output is smooth and uniform; however, with higher order gains such as K/s^2, the operator's output is discrete and impulsive. Again, for ease of modeling, the task dynamics is assumed to be K/s or K/s^2.

5.5 The Haptic-Controlled Human-Excavator Model

Recall Figure 5.12, given this as the closed-loop system of operator in excavation task, and applying the general transfer function formula for a closed-loop system,

Forward loop transfer function = $G(_{OL}) = a$ (5.22)

Closed loop transfer function $= G_{(CL)} = \frac{a}{1+a}$ (5.23)

However, the forward transfer function of the operator-excavator model is given by

$$a = G_{CUES}(S)*G_N(S)*G_F(S)*G_{CN}(S)*G_T(S) \qquad (5.24)$$

where $G_{CUES}(S)$ = the transfer function representing cue/signal processing

$G_N(S)$ = the transfer function of the neuromuscular system

$G_F(S)$ = the transfer function of the force generator system

$G_{CN}(S)$ = the transfer function of the central nervous system

$G_T(S)$ = the transfer function of the task dynamics system

From Equation (5.19), the transfer function for integration of visual, auditory and haptic cues can be represented by the equation below,

$$G_{CUES}(S) = \rho_v Y_{v+} \rho_a Y_{a\,+}\, \rho_h Y_h = \rho v e^{-\tau_v s} + \rho a\, e^{-\tau_a s} + \rho h \left(\frac{e^{-\tau_h s}}{0.6s+1}\right) \qquad (5.25)$$

where

$Y_a = e^{-\tau_a s}$,

$Y_v = e^{-\tau_v s}$,

$Y_h = \dfrac{e^{-\tau_h s}}{0.6s+1}$ and

$\rho_v + \rho_a + \rho_h = 1$

The neuromuscular dynamics has a transfer function given by Equation 5.21 (McRuer, 1980) as

$$G_N(S) = \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2 + 2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} \qquad (5.26)$$

The transfer function for the central nervous system for this model is assumed to be proportional given by equation 5.20*b* (Hess, 1985).

$$G_{CN}(S) = K_p \qquad (5.27)$$

The transfer function for the task dynamics is model as a second order gain Hess (1985) given by

$$G_T(S) = 1/s^2 \qquad (5.28)$$

Finally the force generator for this model $G_F(S)$ is assumed to be 1.

$$G_F(S) \;\; = 1 \qquad (5.29)$$

Substituting Equations 5.25 through 5.29 into Equation 5.24, gives the forward loop transfer function of the haptic-controlled excavator model as

$$G_{H(OL)}=\left\{\left[\rho v e^{-\tau_v s} + \rho a\, e^{-\tau_a s} + \rho h\left(\frac{e^{-\tau_h s}}{0.6s+1}\right)\right] * \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2+2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} * K_p * \frac{1}{s^2} * 1\right\}$$

(5.30)

Using the general closed-loop transfer function formulation, and substituting Equation 5.30 into Equation 5.23, gives the closed-loop transfer function of the haptic-controlled excavator model $G_{H(LC)}$ as

$G_{H(CL)}=$

$$\frac{\left\{\left[\rho v e^{-\tau_v s} + \rho a\, e^{-\tau_a s} + \rho h\left(\frac{e^{-\tau_h s}}{0.6s+1}\right)\right] * \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2+2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} * K_p * \frac{1}{s^2} * 1\right\}}{\left[1+\left\{\left[\rho v e^{-\tau_v s} + \rho a\, e^{-\tau_a s} + \rho h\left(\frac{e^{-\tau_h s}}{0.6s+1}\right)\right] * \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2+2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} * K_p * \frac{1}{s^2} * 1\right\}\right]}$$

(5.31)

Simplifying equation 5.31 gives

$$G_{H(CL)} = \frac{K_H K_n K_p e^{-\tau_n s}}{(Tns+1)\left[\left({s}/{\omega_n}\right)^2+2\left({\xi_n}/{\omega_n}\right)s+1\right]s^2+K_w K_n K_p e^{-\tau_n s}} \quad (5.32)$$

where $K_H = \rho v e^{-\tau_v s} + \rho a\, e^{-\tau_a s} + \rho h\left(\frac{e^{-\tau_h s}}{0.6s+1}\right)$

5.6 Traditional Human-Excavator Model

In the traditional human-excavator model, the operator relies mainly on visual and auditory cues in order to perform a given excavation task. The instantaneous exchange of information between operator and machine that is associated with haptic modality is absent. A representation of the traditional human-excavator model is shown in Figure 5.15. In this model, the operator manipulates the excavator by using levers, pedals, joysticks as the input control devices.

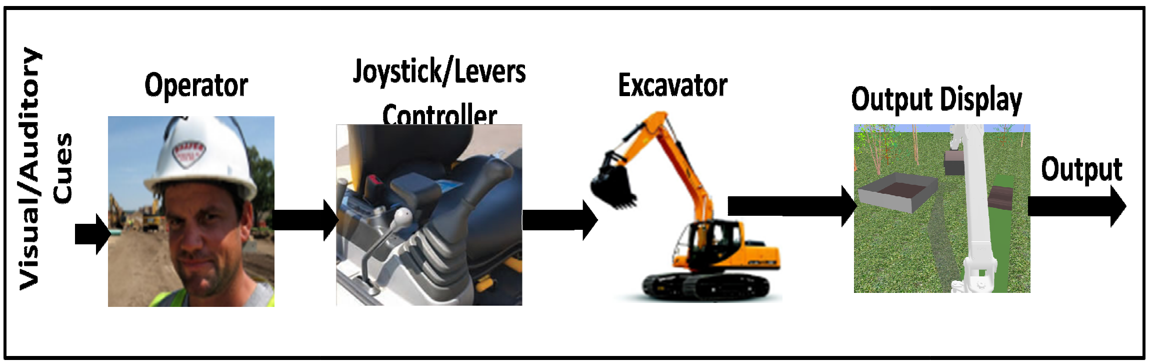

Figure 5.15: Traditional human-excavator model

The structural representation of the traditional human-excavator model is shown in Figure 5.16. The visual and auditory cues with known transfer functions Y_v, and Y_a are integrated based on how a particular cue contributes to overall perception. For a bimodal information source with visual and auditory cues, McRuer (1980), Hess (1985), Deng (1999), the transfer function of the cue integration is given by

$$G_{CUES}(S) = \rho_v Y_{v+} \rho_a Y_a = \rho_v * e^{-\tau_v S} + \rho_a * e^{-\tau_a S} \qquad (5.33)$$

where $Y_{v=} \mathrm{v} e^{-\tau_v S}$ is the transfer function of the visual cue,

$Y_{a=}$ v$e^{-\tau_a s}$ is the transfer function of auditory cue, ρv is the probability of occurrence of visual cue, ρa is the probability of occurrence of auditory cues, and

$\rho_v + \rho_a = 1$

The structural model of the bimodal visual-auditory traditional excavator is shown Figure 5.16 below.

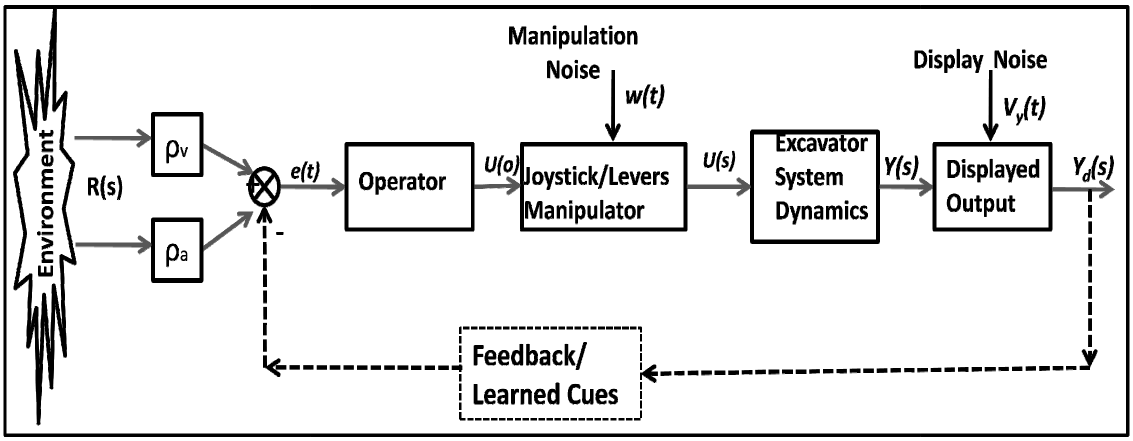

Figure 5.16: Structural representation of traditional human-excavator model

Given the forward-loop transfer function of the operator-excavator system, *a,* is given by $G_{T\,(OL)} = a$, and the closed loop transfer function is given by $G_{T\,(CL)} = a/1+a$,

$$G_{T\,(OL)} = a = G_{CUES}(S)*G_N(S)*G_F(S)*G_{CN}(S)*G_T(S) \qquad (5.34)$$

where $G_{CUES}(S)$ = the transfer function of cues present in environment

$G_N(S)$ = the transfer function of the neuromuscular system

$G_F(S)$ = the transfer function of the force generator system

$G_{CN}(S)$ = the transfer function of the central nervous system

$G_T(S)$ = the transfer function of the task dynamics system, then substituting

$G_{CUES}(S) = \rho_v * e^{-\tau_v s} + \rho_a * e^{-\tau_a s}$, and Equations *(5.26)-(5.29)* into Equation *(5.34)* gives the forward loop transfer function of the traditional excavator model

$$G_T(OL) = \left\{ \left[\rho_v * e^{-\tau_v s} + \rho_a * e^{-\tau_a s}\right] * \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2 + 2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} * K_p * \frac{1}{s^2} * 1 \right\} \tag{5.35}$$

Therefore, the closed loop transfer function of the traditional excavator model is given by

$$G_T(CL) = \frac{\left\{ \left[\rho_v * e^{-\tau_v s} + \rho_a * e^{-\tau_a s}\right] * \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2 + 2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} * K_p * \frac{1}{s^2} * 1 \right\}}{\left(1 + \left\{ \left[\rho_v * e^{-\tau_v s} + \rho_a * e^{-\tau_a s}\right] * \frac{K_n e^{-T_n s}}{(T_n S+1)\left[\left(\frac{S}{\omega_n}\right)^2 + 2\left(\frac{\xi_n}{\omega_n}\right)S+1\right]} * K_p * \frac{1}{s^2} * 1 \right\}\right)} \tag{5.36}$$

Simplifying Equation (5.36) gives

$$G_T(CL) = \frac{K_T K_n K_p e^{-\tau_n s}}{(TnS+1)\left[\left(S/\omega_n\right)^2 + 2\left(\xi_n/\omega_n\right)S+1\right]S^2 + K_T K_n K_p e^{-\tau_n s}} \tag{5.37}$$

where $K_T = \left[\rho_v * e^{-\tau_v s} + \rho_a * e^{-\tau_a s}\right]$

5.7 Model Representation in Matlab

In order to investigate and compare the characteristics of the haptic-controlled excavator model with the traditional excavator model, both were implemented and analyzed as control systems using Matlab simulation software. A major characteristic of a control system is its stability which can be determined from the Bode plots and Nyquist diagrams. A control system is said to be stable if its impulse response approaches zero (0) as time approaches infinity or if every bounded input produces a bounded output. In other words, a control system is said to be stable if its natural response decays to zero with time. By implementing in Matlab, Bode plots and Nyquist diagrams of both systems were obtained and compared. A major characteristic of Matlab is that it allows different characteristics of the models such as Bode plots and Nyquist diagrams of the two systems to be compared.

In order to implement the models, the two model equations were modified, simplified and represented as polynomial functions. First, the forward-loop function of the haptic-controlled human-excavator model Equation 5.32 was simplified and modified into

$$G_{H(OL)} = \frac{K_H K_n K_p e^{-T_n s}}{\left({}^{T_n}/_{\omega^2}\right)s^5+\left(\frac{2T_n\xi}{\omega}+\frac{1}{\omega^2}\right)s^4+\left(T_n+\frac{2\xi n}{\omega_n}\right)s^3+s^2} \qquad (5.38)$$

where $K_H = \rho v e^{-\tau_v s} + \rho a\, e^{-\tau_a s} + \rho h\left(\frac{e^{-\tau_h s}}{0.6s+1}\right)$

and closed-loop transfer function $G_{H(CL)}$ is given by

$$G_{H(CL)} = \frac{\left({}^{T_n}/_{\omega^2}\right)s^5+\left(\frac{2T_n\xi}{\omega}+\frac{1}{\omega^2}\right)s^4+\left(T_n+\frac{2\xi n}{\omega_n}\right)s^3+s^2}{\left({}^{T_n}/_{\omega^2}\right)s^5+\left(\frac{2T_n\xi}{\omega}+\frac{1}{\omega^2}\right)s^4+\left(T_n+\frac{2\xi n}{\omega_n}\right)s^3+s^2+0s+K_H K_n K_p e^{-T_n s}} \qquad (5.39)$$

The Bode plot and Nyquist diagram of the model is shown in Figure 5.17 and Figure 5.18 respectively. In both figures, the blue line represents the open-loop transfer function; the green line represents the closed-loop transfer function, and the circled red-cross represents the -1 position on the real axis of the Nyquist plot.

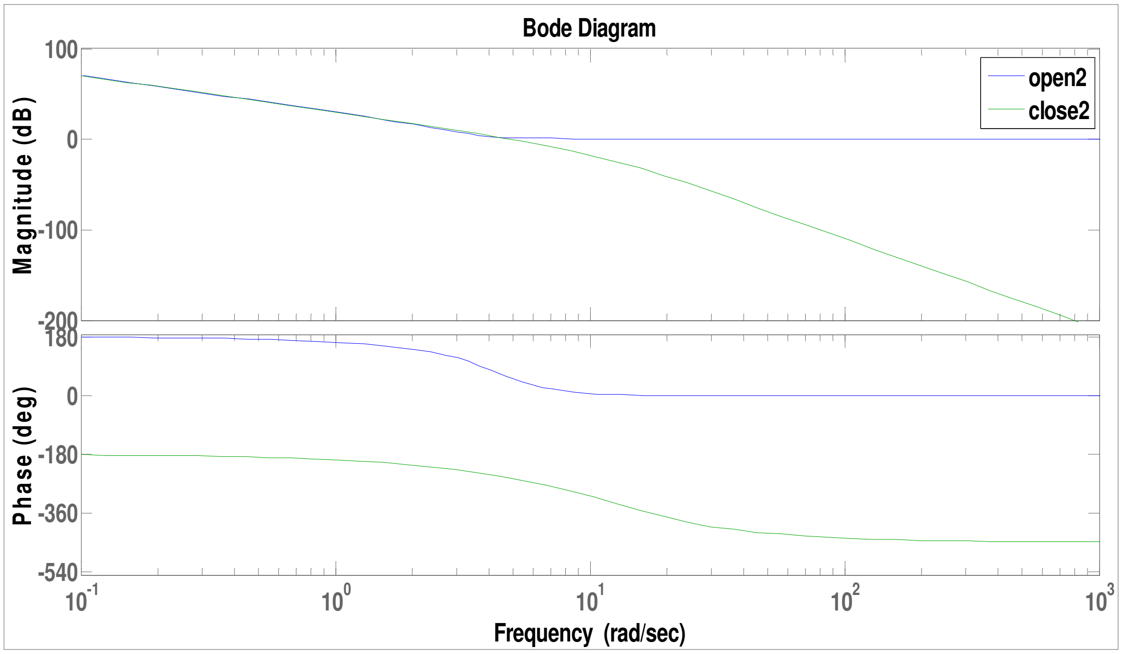

Figure 5.17: Bode plot of haptic-controlled human-excavator model

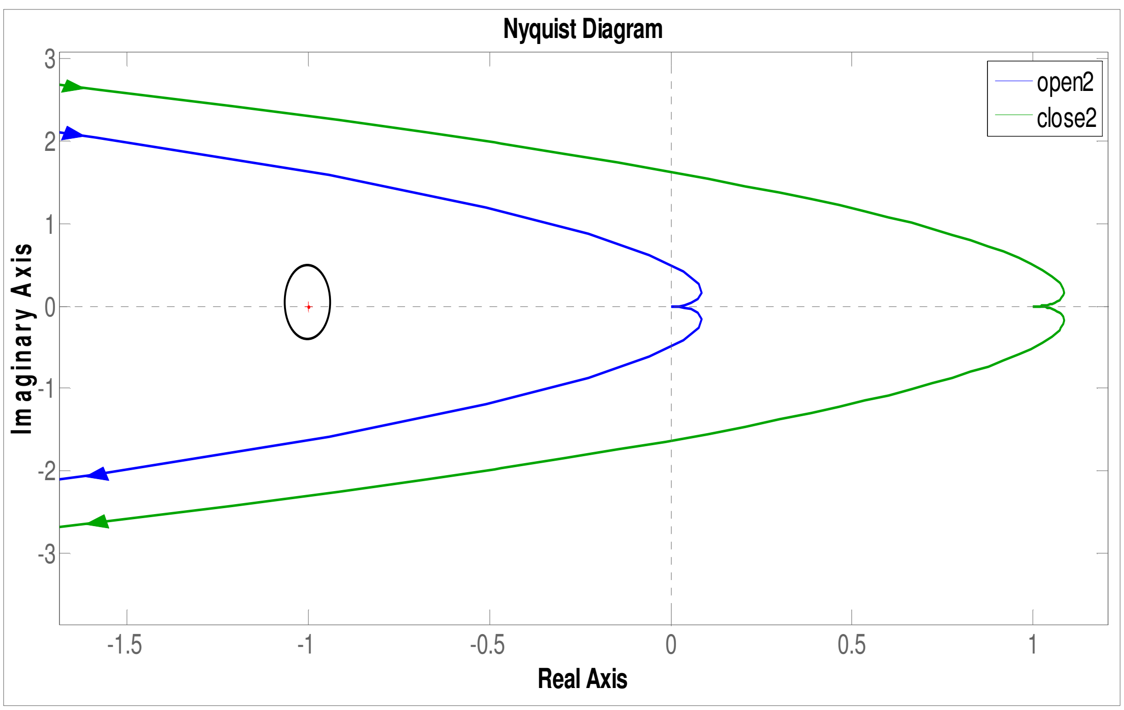

Figure 5.18: Nyquist plot for the haptic-controlled human-excavator model

Similarly, for the traditional human-excavator model, the modified and simplified forward loop transfer function $G_T(_{OL})$ is

$$G_T(_{OL}) = \frac{K_T K_n K_p e^{-T_n s}}{\left({}^{T_n}/_{\omega^2}\right)s^5 + \left(\frac{2T_n\xi}{\omega} + \frac{1}{\omega^2}\right)s^4 + \left(T_n + \frac{2\xi n}{\omega_n}\right)s^3 + s} \tag{5.40}$$

and the closed loop transfer function $G_T(_{CL})$ is given by

$$G_T(_{CL}) = \frac{\left({}^{T_n}/_{\omega^2}\right)s^5 + \left(\frac{2T_n\xi}{\omega} + \frac{1}{\omega^2}\right)s^4 + \left(T_n + \frac{2\xi n}{\omega_n}\right)s^3 + s^2}{\left({}^{T_n}/_{\omega^2}\right)s^5 + \left(\frac{2T_n\xi}{\omega} + \frac{1}{\omega^2}\right)s^4 + \left(T_n + \frac{2\xi n}{\omega_n}\right)s^3 + s^2 + 0s + K_T K_n K_p e^{-T_n s}} \tag{5.41}$$

The Bode plot and Nyquist diagram of the model is shown in Figure 5.19 and Figure 5.20 respectively. In these figures, red line represents the open-loop transfer function, while the black line represents the closed-loop transfer function. The red-cross represents the -1 position on the real axis of the Nyquist plot.

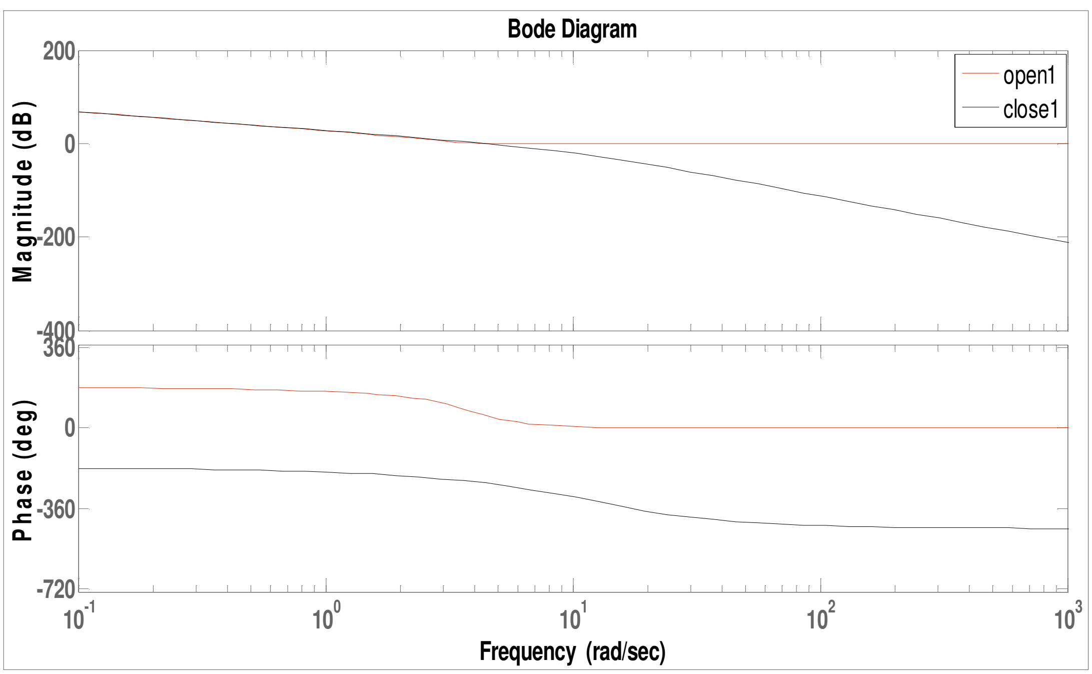

Figure 5.19: Bode plot of traditional human-excavator model

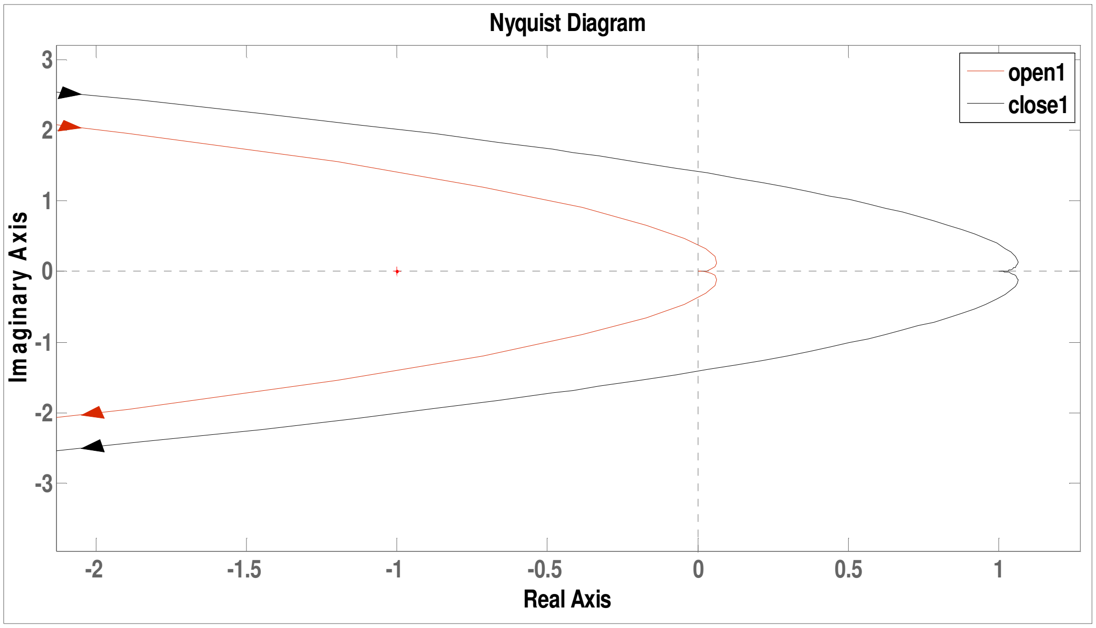

Figure 5.20: Nyquist plot for traditional human-excavator model

From Figure 5.17, the Bode plot for the closed-loop transfer function of the haptic-controlled excavator model is observed to lie below zero indicating stability. The corresponding Nyquist diagram show that the Nyquist plot does not encircle (-1, 0) position as indicated by the '*red-cross*' in Figure 5.18, also an indication that the haptic-controlled excavator model is stable. Also, from Figure 5.19, the Bode plot for the closed-loop transfer function of the traditional excavator model is observed to lie below zero indicating that the model is stable, and the corresponding Nyquist plot of the traditional excavator model does not encircle -1 (indicated by '*red-cross*'), a further indication of the model's stability as shown in Figure 5.20. Results from model implementation shown in Figures 5.17-5.20, therefore, suggest that both haptic-controlled excavator and the traditional excavator models are stable systems.

To compare the stability of haptic-controlled excavator model and the traditional excavator model, both models were plotted on a single bode diagram as shown in Figure 5.21. From the plot, it can be observed that both the haptic-controlled excavator model and the traditional excavator models are stable, as the Bode plots of their closed-loop transfer plots lie below zero. However, the haptic-controlled excavator model appears more stable than the traditional excavator model because the Bode plot of the closed-loop haptic-controlled excavator model (green line) lie below the Bode plot for the closed-loop traditional excavator model (red line) form plot.

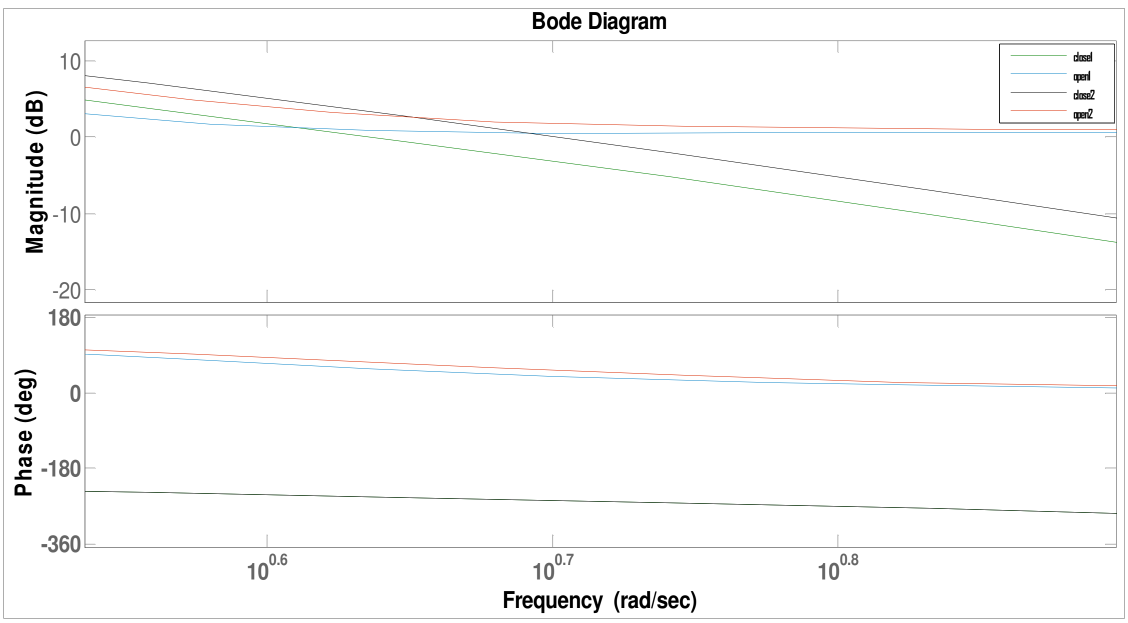

Figure 5.21: Bode plot for haptic-controlled and traditional excavator models

The Bode plots for the open-loop transfer functions of the haptic-controlled excavator model and the traditional excavator model are represented by 'black' and 'blue' lines respectively in Figure 5.21. The results from the modeling exercise show that though both the traditional excavator and the haptic-controlled excavator models are stable, the haptic-controlled excavator model is more stable. This means that operators will be able to use the haptic-controlled excavator more efficiently and effectively than the traditional excavator model. This may probably be due to that fact that the additional communication channel provides additional resources that helps the operator in executing the excavation task.

Further, the Bode plots of visual only, auditory only and haptic only sensory information processing within the haptic-controlled excavator interface were plotted and compared, Figure 5.22. Blue represents auditory only, pink represents haptic only, red represents visual only processing.

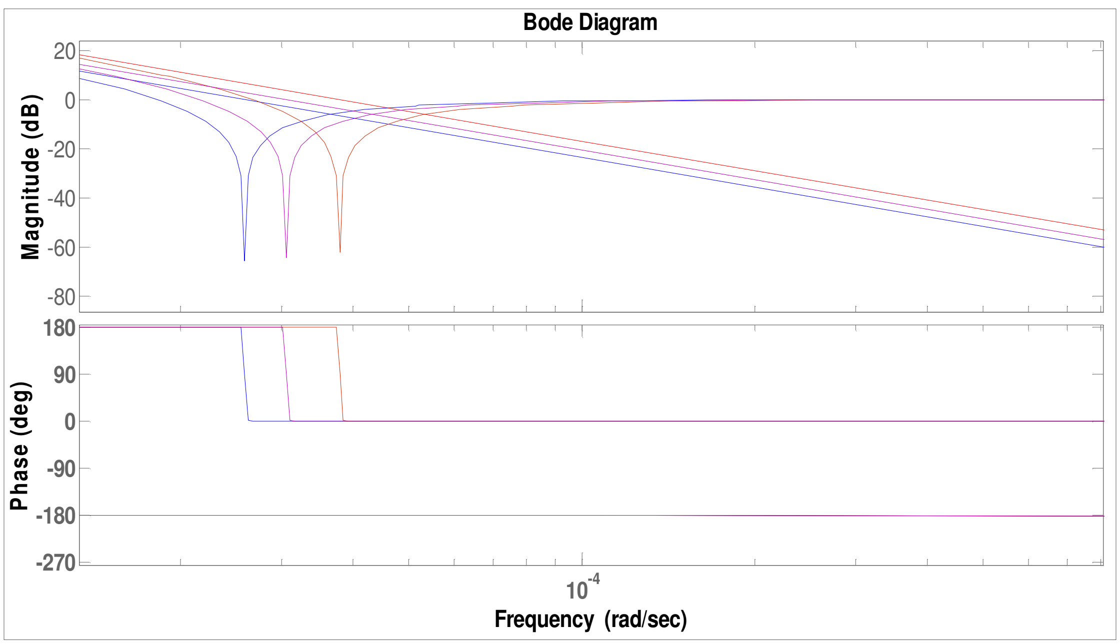

Figure 5.22: Bode plot for visual, auditory and haptic sensory processing

The plot showed that visual only information processing had the highest gain and phase shift angle, followed by the haptic processing, while auditory processing gain the least in terms of magnitude and phase shift. Also, the processing of visual-auditory, visual-haptic, and auditory-haptic multisensory information processing within the domain of haptic-controlled excavator interface was developed and compared as shown in Figure 5.23. Visual-auditory, visual-haptic and auditory-haptic multisensory information processing are represented by green, blue and yellow lines respectively. Again, it can be observed from the plot that visual-haptic multisensory processing had the highest gain and phase shift, followed by visual-auditory processing, while auditory-haptic processing gained the least in magnitude and phase shift.

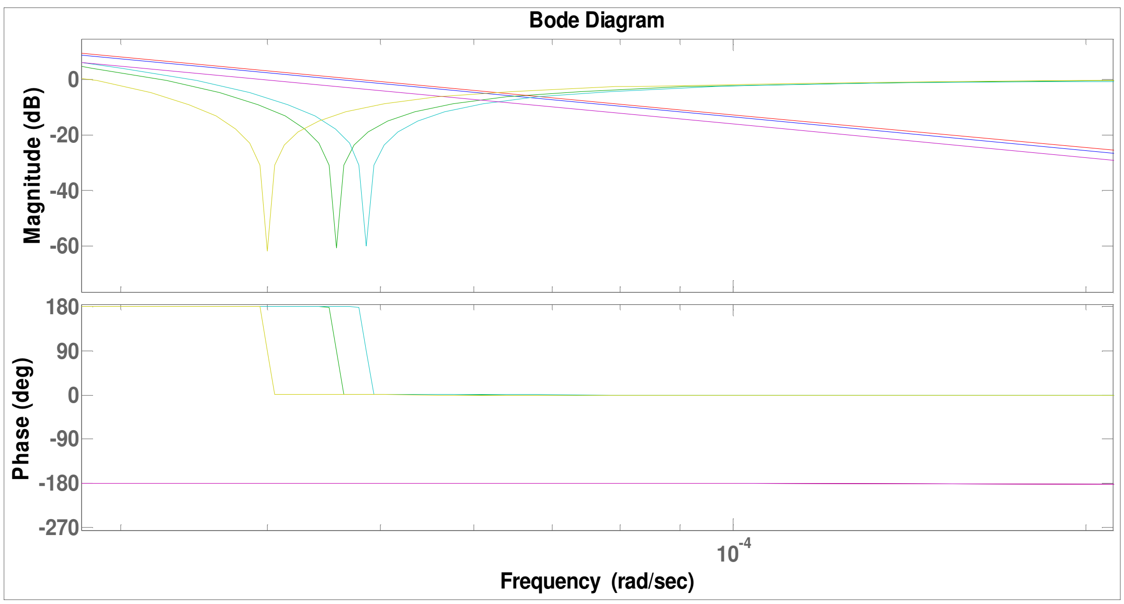

Figure 5.23: Bode plot for visual-auditory, visual-haptic, and auditory-haptic multisensory processing

Finally, visual-auditory, visual-haptic and auditory-haptic multisensory information processing were compared to visual-auditory-haptic multisensory processing as shown in Figure 5.24. As expected, the visual-auditory-haptic multisensory processing (represented by black line) had the highest gain and phase angle shift.

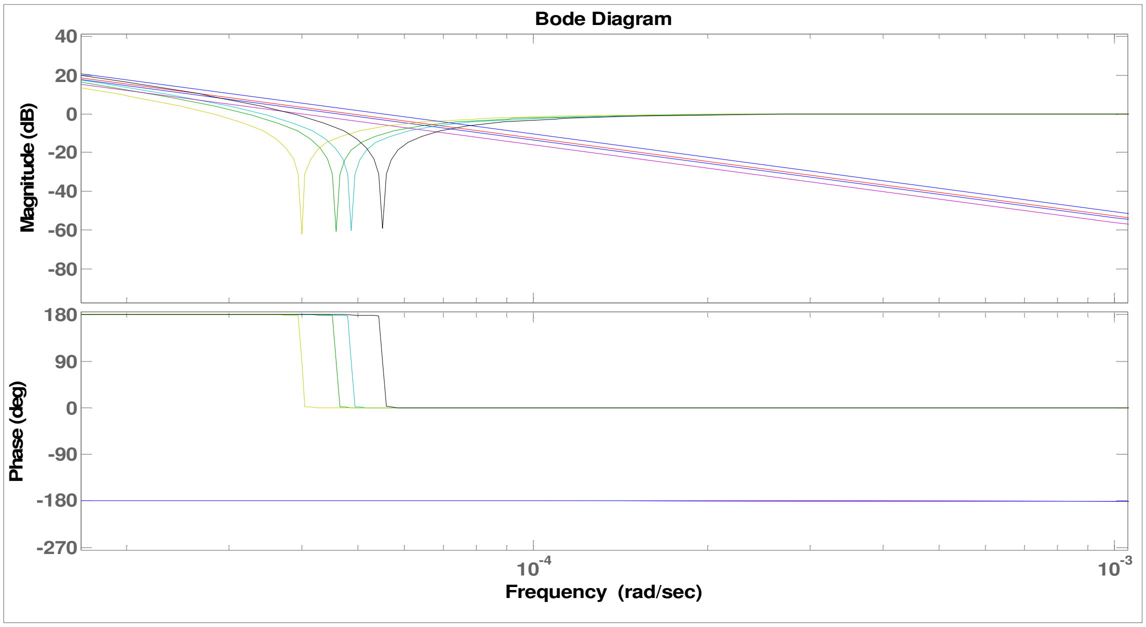

Figure 5.24: Bode plot for visual-auditory, visual-haptic, auditory-haptic, and visual-auditory-haptic multisensory processing

5.8 Chapter Summary

This chapter provided a brief overview of control theory and its application in modeling human information processing. A conceptual model of human-excavator interaction was developed, and a representation of the components of the human-excavator system was presented. Structural model of the human-excavator system was proposed, a transfer function for each component of the structural model was developed for both haptic-controlled excavator and traditional excavator models. The haptic-controlled excavator model and the traditional excavator model were implemented in Matlab and compared using Bode plots and Nyquist diagrams. The results showed that, both haptic-controlled and traditional excavator models are stable systems, however, the haptic-controlled excavator model was more stable and, therefore, easier to control by the operator.

CHAPTER 6

USING USER-CENTERED DESIGN TO IMPROVE HAPTIC-CONTROLLED EXCAVATOR INTERFACE

One of common problems found in many fluid power system designs is that, engineers often are concerned with the utility of the system (i.e. whether the functionality of the system in principle can do what it is supposed to do) and, therefore, pay little attention to its usability (i.e. how well users can use the system to accomplish a given task). Often times, this lead to systems with high functionality but not as user friendly. As a result, operators have to be trained for long periods of time, in order to learn how to use these machines. Further, operators sometimes have to operate machines in an uncomfortable posture/position for long periods, which sometimes lead to cumulative trauma disorders (CTDs). The cost associated with long operator training and medical treatment for cumulative trauma disorders among other costs can have a major financial impact on the bottom-line of companies that use fluid power systems.

One design strategy that is often employed in designing systems that are user friendly is the User-Centered Design (UCD) approach which broadly describes design process in which end-users influence how a design takes shape. The concept of user-centered design was first coined by Norman and Draper (1986) to emphasize the importance of having a clear understanding of the users but without necessarily involving them actively in the design (Gulliksen et al., 2003).

With this design approach, the active involvement of users is sought in order to understand clearly user needs and requirements. This is done iteratively throughout the design and evaluation process usually involving a multi-disciplinary team working together to achieve the desired results. The goal is to design interface for fluid power systems that are intuitive, efficient, easy to learn and use, and help operators accomplish tasks while at the same time avoiding mistakes.

It is not uncommon to come across everyday products/objects with poor and unintuitive designs and often leave users frustrated and unable to perform simple tasks. Most people have had the experience where they bought a product or gadget, took it home and found themselves frustrated and unable to use the product because the design is unfriendly, a problem often compounded by instructions that are difficult to understand. Similarly, it has been documented that excavator operators often need to be trained for long periods of time, in order to ensure that, they learn the inverse kinematic relationships between the lever displacement and bucket trajectory (Frankel, 2004).

User-centered design is an approach to design that seeks to develop products that are more usable and support users do their tasks by involving users throughout the design process. The user-centered design is loosely defined as a method for designing ease of use into the total user experience, through improved usability of product/system by placing the needs of the intended end users at the core of product design (Norman & Draper, 1986). The term user-centered design originated in Norman's research laboratory at the University of California San Diego (UCSD) in the 1980s and became widely used after he co-authored the book 'User-Centered System Design: New Perspectives on Human-Computer Interaction' with Draper in 1986 (Abras et al., 2004). User-centered design emphasizes that the purpose of the design is to serve the user, not to use the specific technology, nor to be an elegant piece of programming, but rather, the needs of users should dominate design of the interface, and the needs of the interface should dominate the design of the rest of the system (Norman & Draper, 1986).

User-centered design focuses on the requirements of potential users from the product's inception, and checks at each step of the design phase with users to ensure ease of use as well as user satisfaction with the final interface/product design (Norman, 1988). The user-centered design process puts user needs at the center of the design, involves users throughout the all phases of the design. User-centered design has been proven to be an effective design strategy that ensures ease of use, safety, and effectiveness of the interface.

The International Standards Organization (ISO) established standards in 1999, ISO 13407, provides guidance on user-centered design activities throughout the design and life-cycle of computer-based interactive systems in order to manage the design process, and describe user-centered design from four different perspectives: namely the rationale, principles, planning and activities of the user-centered design (Jokela et al., 2003).

Rationale: The rationale for using user-centered design in developing the haptic-controlled excavator interface, as has been mention elsewhere, is to ensure intuitive, safe, effective, easy to use interface that is responsive to operator commands and, therefore, improves user satisfaction and productivity while reducing costs associated with operator training.

Principles: The four general principles (active involvement of users and clear understanding of user and task requirements, appropriately allocating resource functions between users and technology, continuous iteration of design improvements, and working with multi-disciplinary team to ensure input from diverse background) established by ISO 13407 for user-centered design was followed.

Planning: Activities of work were planned and conducted so that results and recommendations could be used to modify the haptic-controlled excavator interface.

Activities: These activities describe the core of the user-centered design effort. Figure 6.1 shows a representation of activities in each phase of the user-centered design process.

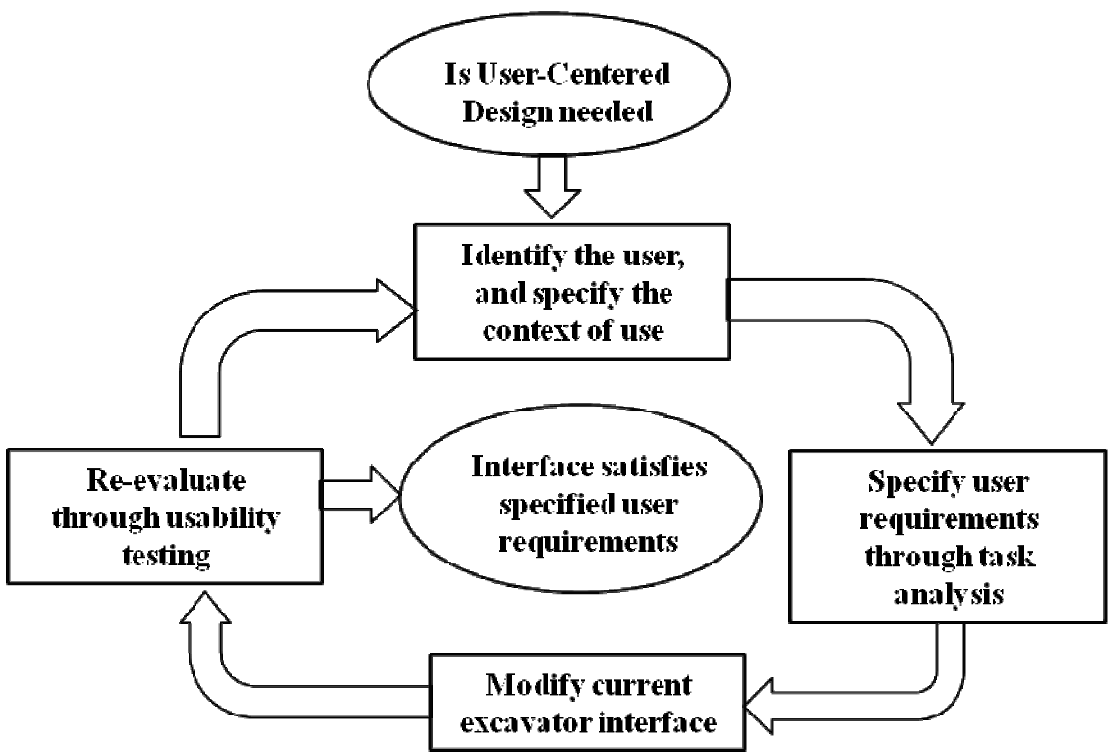

Figure 6.1: Activities of user-centered design approach

Typically, the user-centered design process is an iterative process, and involves a usability evaluation of the current interface/product (or in the case of a new product, evaluation of similar products), to identify user requirements and needs through interviews and observations, conducting task analysis, setting up usability goals, developing prototypes, and conducting usability testing (Vredenburg et al., 2001). The following describes how each phase of the user-centered design process for the haptic-controlled excavator interface was achieved.

1. First, a user profile is developed to provide information to designers about whom they designing the product for. This is accomplished by conducting interviews and surveys with potential users to identify their characteristics, needs and requirements.
2. Second, in order to design products that meet the task needs of the operators, task analysis was carried out to gain good understanding of the nature of excavation tasks and how they are performed. A hierarchical task analysis was conducted to understand common excavation tasks through direct observation and interviews. From the results of the task analysis, usability goals for the haptic-controlled excavator interface were determined.
3. Upon completion of task analysis, a usability testing was conducted using the haptic-controlled excavator interface currently under construction at the Georgia Institute of Technology, in Atlanta. The goal was to help identify potential usability problems in the interface and provide recommendations for improvement. Further, two empirical investigations were conducted. First, to

assess conflicts in the haptic-controlled excavator interface and the impact such conflicts might have on the performance of operators, and second, to determine the range of force feedback values that produce best operator performance.

4. Results and recommendations from the design phases above were used to modify and improve the current haptic-controlled excavator interface.

Using the user-centered design process described above ensures that it is easy for operators to determine what actions are possible while interacting with the haptic-controlled excavator. It ensures operators can easily evaluate the current state of the system, and help them follow the natural mapping between their intentions and actions, and between actions and resulting effect, as well as between visible information and the interpretation of the system state (Abras et al., 2004; Norman, 1988). The following sections describe each of the user-centered design process.

6.1 User Profile

To help construct user profile for excavator operators and to identify user characteristics, needs and requirements, six excavator operators were interviewed. A summary of user profile and characteristics described next. First, all participants interviewed were males between the ages of 37 to 54 years old, with average age of about 50 years. The number of years of excavator operating experience ranged from a minimum of 6 years experience to 27 years. When asked about their computer literacy, four participants said they considered themselves moderate to expert computer users, while two considered themselves as novices in terms of computer use. Majority said they use computer to surf the internet (Facebook, Craigslist), send and receive emails, play online games. When asked if they would learn new computer skill, majority said they would, however, there was one participant who said he was not enthusiastic about learning new computer skill, though he added that if he needed to learn a new computer skill in order to perform his job, then he was willing to learn the new skill.

All participants normally worked 8 hours/day for 5 days/week, all said they worked for extra hours/day and extra days/ week if their services were needed by their employer. When asked to describe the tasks that they perform in a typical day on the job, participants described machine operation (digging, scooping, loading, unloading, leveling, filling, piling, moving), as the primary tasks, together with other tasks such as preventive maintenance, problem diagnosis, ability to read grade plans and use grade stakes to measure the amount of earth removed, follow both spoken and hand signals.

When participants were asked if they experienced fatigue while performing excavation tasks, five said they sometimes experienced fatigue and attributed such fatigue to cabin vibration, long periods of sitting, and the sometimes unfriendly work environment in which excavation task is performed. One participant, however, said he rarely experienced any fatigue and attributed this to his physical and mental strength, which according to him, helps him withstand the sometimes harsh environment in which he works. When probed specifically about shoulder and wrist fatigue, most participants said they experienced wrist fatigue from time to time due to twist and turn motions of the wrist when operating the excavator, however, only two participants felt shoulder fatigue was an issue to them, while the rest did not see shoulder fatigue as big problem.

6.2 Task Analysis

Task analysis generally describes the physical tasks and cognitive plans required of a user to accomplish a particular goal. It includes a detailed description of both manual and mental activities, task durations and frequency, task allocation and complexity, environmental conditions, and any other unique factors involved in or required for user to perform a given task (Hone & Stanton, 2007; Kirwan & Ainsworth, 1992; Stanton, 2006). It is the fundamental methodology used in the assessment and reduction of human error. Task analysis is used to analyze tasks that users of a system are expected to perform in order to eliminate the preconditions that give rise to errors before they occur, and can be used to aid in the design stage of a new system, the modification of an existing system, or as part of an audit of an existing system (Embrey, 2000).

Several task analysis methods exist; however only hierarchical task analysis and cognitive task analysis are discussed in this work. A hierarchical task analysis is a systematic method of describing how work is organized in order to meet the overall objective of job and involves identifying in a top down fashion the overall goal of the task, then the various sub-tasks and the conditions under which they should be carried out to achieve that goal (Embrey, 2000).

A hierarchical task analysis is a graphical representation of the decomposition of the high level tasks into constituent subtasks, operations, and actions used to accomplish them. Hierarchical task analysis can be described by goals, tasks and actions. The goals define what a user wishes to achieve, a task represent one of the activities that must be performed in order to achieve the goal, and an action is a simple task which has no further structure or the lowest level of decomposition (Hone & Stanton, 2007). Cognitive task analysis on the other hand, models the internal representation and processing that users follow to perform a task. Cognitive task analysis is a methods use set of tools, techniques and protocols to identify the cognitive skills and mental demands needed to perform a task efficiently (Prasanna, Yang, & King, 2009).

For this work, a hierarchical task analysis was used to help decompose excavator operation tasks and identify the crucial task necessary for successful completion of excavation task.

6.2.1 Task Analysis of Haptic-Controlled Excavator Interface.

The goal of conducting task analysis is to understand the critical tasks that excavator operators perform, so that operator-excavator interface is designed to assist operators to carry out these tasks. Task analysis is the breakdown of how a task is accomplished, including a detailed description of both manual and mental activities, task and element durations, task frequency, task allocation, task complexity, environmental conditions, necessary clothing and equipment, and any other unique factors involved in or required for one or more people to perform a given task. A task analysis breaks the excavation tasks into goals, tasks and actions needed to complete them successfully. The goals are what the operator wishes to achieve, tasks are the activities which must be performed in order to achieve those goals, and actions are the simple tasks (with no further structure) that must be performed to accomplish a task. A high level flow chart of the excavation task is shown in Figure 6.2.

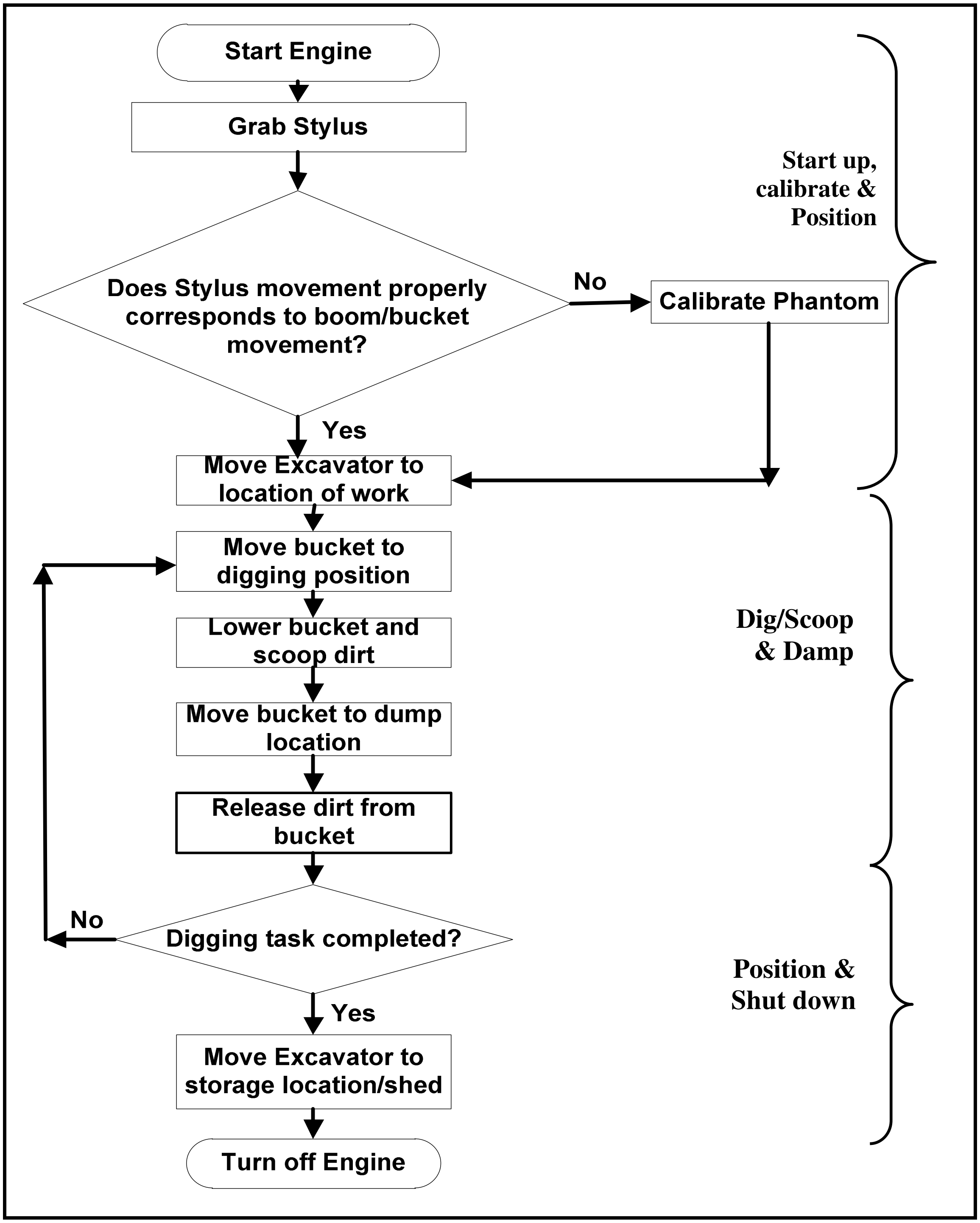

Figure 6.2: Flow chart of excavation task

In order to understand the tasks that excavator operators perform, excavator operators were observed as they performed their tasks and then interviewed to understand the operators' mental processes as they control and manipulate excavators. A hierarchical task analysis was conducted to break the excavation task into goals, sub goals, tasks and actions (simple operations) required to accomplish a given excavation task. The hierarchical task analysis result for the excavation task is shown in Figure 6.3.

Also, the Goals, Operators, Methods and Selection (GOMS) method of task analysis (Card, Moran, & Newell, 1980) was used to decompose the excavation tasks into component operations for thorough understanding of task steps (Appendix C). Table 6.1 shows typical tasks performed in an excavation task and their relative importance. From the task analysis, operator interviews and observations, the following tasks were identified as the most critical tasks necessary for successful completion of a given excavation task.

1) Move boom/bucket
2) Scoop/dig dirt
3) Position boom/bucket
4) Open/close bucket
5) Load/unload bucket content

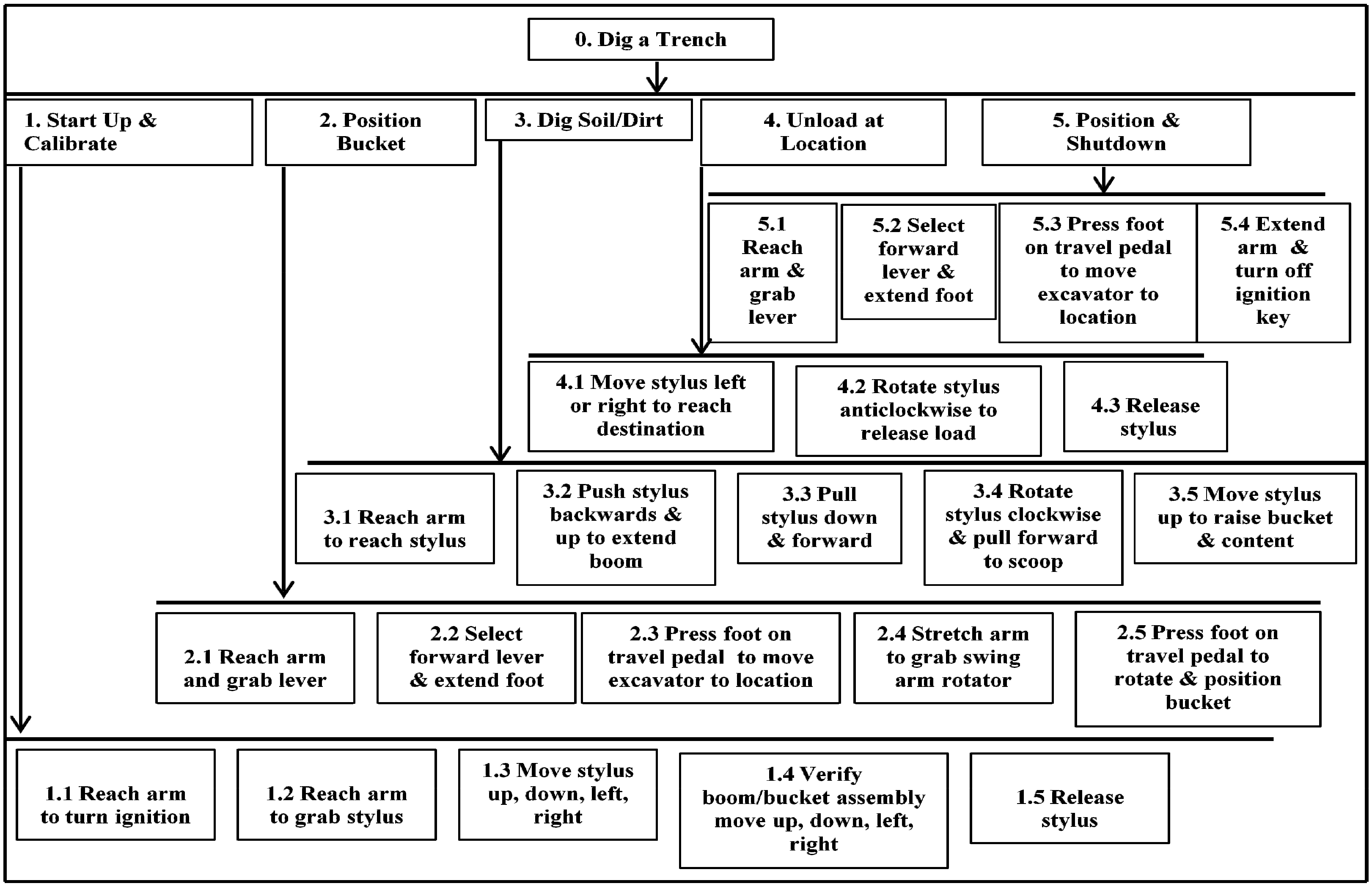

Figure 6.3: HTA for excavation task using haptic control excavator

Table 6.1: Tasks in typical excavation task and their relative importance

Task	Importance
Start excavator	High
Calibrate haptic device	Medium
Move to excavator to work location	Low
Position boom & bucket	High
Scoop/dig dirt	High
Move bucket to damp site	High
Unload content of bucket	High
Move to storage	Medium
Shut down	Medium

6.3 Usability Goals

These usability goals were developed to help focus the attention and resources of the haptic-controlled excavator interface design team on user and issues important to them. This allowed the design team to focus on the ‘voice of users’ throughout all stages by continuously evaluating and testing through user interaction. Both qualitative and quantitative usability goals were developed for the haptic-controlled excavator interface. Qualitative usability goals are summarized in Table 6.2, and identify the critical tasks and design limitations that must be improved to increase the effectiveness and user friendliness of the haptic-controlled excavator interface. The qualitative usability goals were derived from the task analysis as well as the empirical investigations discussed in Chapter 3 and Chapter 4. Further, quantitative usability goals were developed to help define ease of use as well as ease of learning of the haptic-controlled excavator interface. The quantitative usability goals were formulated in terms of performance goals based on the two empirical studies conducted as part of this dissertation.

Table 6.2: Qualitative usability goals for haptic-controlled excavator

Task/design area	Usability goal
Move boom/bucket	The sensitivity of the haptic device (stylus) must be improved to ensure steadiness and firmness of stylus in order to ensure smooth, controlled movement of the boom/bucket assembly.
Scoop/dig dirt	Bucket open/close cycle must correspond appropriately with clockwise/anticlockwise stylus rotation to give operator control while manipulating the excavator. The current design is too sensitive and unsteady making operator control very difficult. This can lead to errors.
Position boom & bucket	Excavator boom/bucket assembly must correspond well with stylus and pen of haptic device to ensure complete controllability and ease of use
Open/close bucket	Opening and closing bucket must accurately correspond to clockwise and anticlockwise rotation of stylus. This will stabilize the controls and make the haptic-controlled excavator easy to learn and use and help reduce errors
Load and unload	Provide steadiness and firmness to stylus to ensure complete operator control by adding proportional weight to content of bucket. Provide weight to ensure sense of realness
Adequate workspace around haptic device	Providing adequate workspace around the haptic device to ensure uninterrupted operator control and safe operation of excavator and prevent errors
Haptic force feedback	Provide appropriate force feedback to allow steadier and effective control by operator as identified in the empirical study

Table 6.2: Qualitative usability goals for haptic-controlled excavator (*cont)*

Task/design area	Usability goal
Adequate view of workspace	Provide adequate and unrestricted view of workspace for operators to see clearly at all times. A bucket mounted camera that sends instant video of work space to operator on a monitor mounted in the cabin.
Placement of haptic device	Position haptic device so that it is easily reachable by operators. This will improve hand-eye coordination as the tendency for operators to look at the haptic device will be minimized.
Provision of proper arm rest	Provide a properly designed and well placed arm rest to provide support to operator while using the haptic device. This reduce arm and shoulder fatigue.
Bucket weight	Develop and incorporate realistic weight of bucket content to give operators a sense of weightiness or gravity to provide steadiness to system
Scooping sound	Incorporate a realistic scooping sound into interface to provide feedback whenever the bucket scoops/digs to help prevent errors

The performance goals quantified actual user performances while using the haptic-controlled excavator interface to perform excavation task. In setting the quantitative usability goals, task completion time, the number of scoops required to fill a bin, the number of drops outside of the bin, as well as the error rate were used as the performance measures. The quantitative usability goals for the haptic-controlled excavator interface are summarized in Figure 6.4. Performance measures for experts and novices were compared. Since this is a new design, performance of novices is used as the current or minimum standard upon which all future iterations will be based. Expert performance is the target performance for the design, while future performance is the performance that the design seeks to achieve in the long term after several iterations.

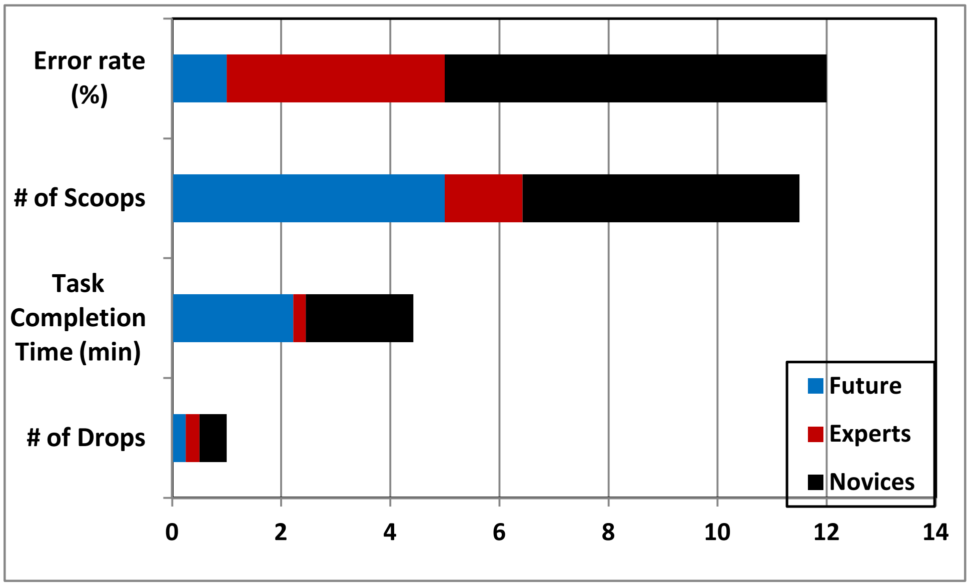

Figure 6.4: Quantitative usability goals for haptic-controlled excavator interface

6.4 Usability Testing

To evaluate the haptic-controlled excavator interface, usability testing was conducted using the prototype haptic-controlled excavator simulator under construction at the Georgia Institute Technology in Atlanta, GA. The purpose of this usability evaluation was to investigate the actions and behaviors of operators as they interact with a haptic-controlled excavator, and identify potential usability problems that may confront operators.

Usability of a system/interface refers to the ease with which users are able to use the interface to accomplish the required task (or a measure of a product's potential to accomplish the goals of the user). According to Nielson 1993, a system's acceptability has two dimensions: practical acceptability and social acceptability. Practical acceptability is defined by usefulness, cost, reliability, compatibility etc. Further, usefulness has two dimensions, utility and usability. Usability can be defined by 5 main attributes. These are *learnability, efficiency, memorability, error rate,* and *satisfaction*. Figure 6.5 represents the attributes of a system's acceptability as well as the different dimensions of usability.

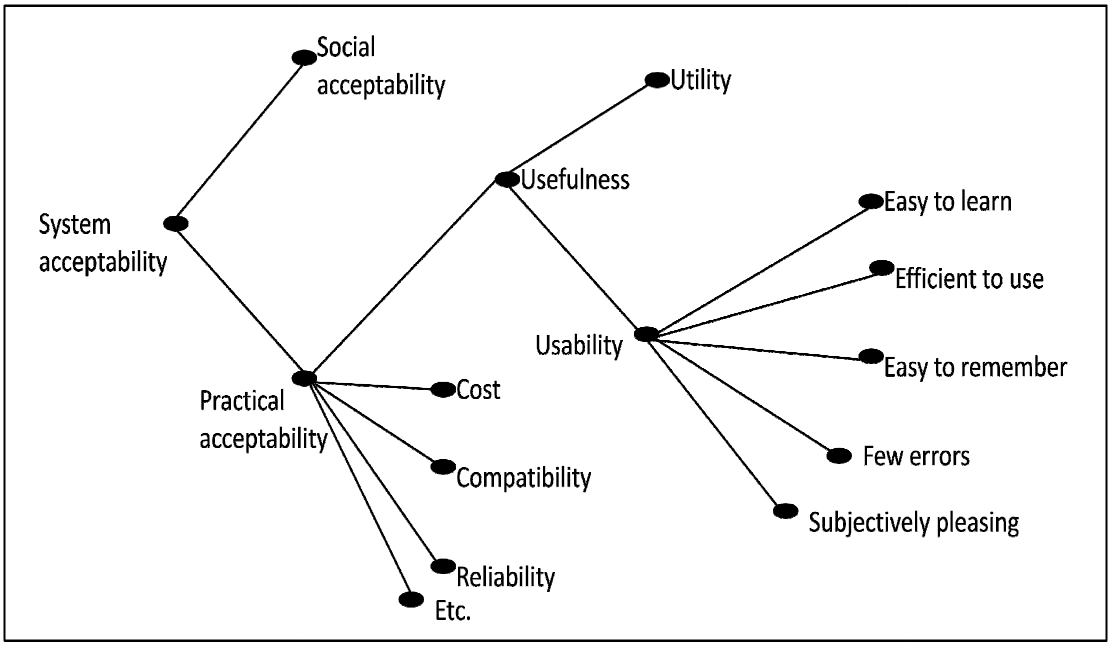

Figure 6.5: Model of attributes of system acceptability [Courtesy of Nielson, 1993]

Usability of a system usually has some tradeoff with utility of the system. While system utility describes whether or not a system performs as designed (system functionality), system usability describes whether or not the user is able to successfully use the system as designed (usefulness). A valid usability test, thus, could yield valuable information similar to what will be expected if the product/interface were to be used outside laboratory settings. To ensure test validity, participants performed tasks that are comparable to the actual tasks performed by excavator operators.

6.4.1 Usability Testing Methodology.

Twenty students, (14 males and 6 females) between the ages 21-31 (mean age = 24, standard deviation = 2.37) were recruited from the Georgia Institute of Technology to take part in the usability testing. The equipment for the experiment consisted of 3 computers, a Bobcat excavator cabin, a Phantom Premium 1.0A haptic device, a 52'' Samsung flat screen LCD and 2 video cameras. The task involved using the stylus of the Phantom Premium device to dig dirt from the marked trench area and dump the dirt into two bins located to the left and right of the trench. The tasks were chosen based on task analysis results that identified moving, digging and dumping/pilling as common tasks often performed by excavator operators.

The test procedure involved briefing participants on the purpose of the study upon arrival, and asking them to read and sign a consent form. A pre-test questionnaire was then administered to collect demographic information. A short demo of the simulation was given, and participants were given a few minutes to familiarize themselves with the simulator. Questions about the simulator and controls from the participants were answered by the experimenter before the test started. All participants were informed that the experiment would be video-taped for further analysis. Upon completion of the tasks, the participants were thanked, debriefed, and asked to complete a post-test questionnaire. Participants were asked about their experience using the haptic-controlled excavator interface, their comfort level, and for their comments and suggestions. Overall, it took about 1 hour to complete the test.

6.4.2 Usability Testing Results.

Results from usability test are summarized in Table 6.3. The usability attributes of *learnability, efficiency, memorability, error rate/prevention* and *user satisfaction* were rated based on participants' responses and other feedback received from participants.

Table 6.3: Usability problems identified, and usability attributes impacted

Usability problem	Usability attribute impacted
1. Excavator too sensitive/stiff to be properly controlled with phantom device/stylus. This may frustrate users and cause fatigue in shoulder and wrist	Efficiency of use, learnability and user satisfaction
2. Operators unable to steadily control excavator. Users found it difficult to maintain control of the excavator with the phantom/device stylus	Efficiency of use, learnability, error prevention
3. Difficulty in maintaining hand-eye coordination due to stiffness and general awkwardness of interface	Efficiency of use, learnability and error prevention
4. Bucket movement is not properly synchronized with rotation of stylus (bucket movement responds poorly to stylus command/rotation)	Memorability, error prevention, user satisfaction
5. No difference between an empty and a full bucket, also no feeling of contact between the bucket and objects (ground, bin, pipe, trench walls, etc)	Learnability of use, memorability and error prevention
6. Difficulty understanding the mapping between excavator and phantom device. Mapping of excavator arm to phantom device is reversed	Learnability of use, memorability error prevention and efficiency of use
7. Lack of appropriate arm rest/support may lead to fatigue in shoulder and elbow	Efficiency of use, user satisfaction
8. Restricted workspace around phantom device may interfere with task performance	Efficiency of use, error prevention, and user satisfaction

On learnability, all participants judged the haptic controlled interface as being easy to learn. Similarly, all participants felt tasks were easy to performed, though some indicated that they found the interface a bit confusing initially. Even with initial confusion, they were able to learn the system fairly easily with little practice. Since nearly all participants were novice users, the design implication is that the haptic controlled interface is generally easy to learn and use, and novice users can learn to use it within a reasonably short period of time with some improvement.

When asked if they were able to efficiently carry out the assigned task using the haptic interface, most participants felt the efficiency of the interface could be improved. First, participants complained that the phantom device was too sensitive and stiff. Either way, it made control of the bucket as well as the movement of boom difficult. About 30% of participants felt that rotating the stylus of the Phantom device did not correspond well enough with open and close movements of the bucket, further most participants reported that the bucket did not respond very well to the rotation command of the stylus, or that the bucket opened/closed while the user had not given any rotation command. The combined effect of the stiffness and the general awkwardness of the phantom control resulted in fatigue and stress in the shoulder and wrist of participants. This prevented users from performing the task in a more efficient manner.

On memorability of the system, most participants felt the interface was easy to remember. It was observed that three participants who have had a previous experience with the haptic interface had average task completion time of 117.13 seconds compared to the overall task completion time of 132.86 seconds for all participants. Clearly, those who had prior experience with the interface were able to complete the tasks faster because they relied on their prior knowledge. A common concern expressed by participants was that the clockwise and counterclockwise movement of the stylus did not correspond well with bucket open/close motion. As a result, users sometimes had to rotate the stylus multiple times in order to open or close the bucket. This led to a situation where participants forgot which direction of rotation corresponded to bucket open or close movement.

On error prevention, most participants felt the high sensitivity and stiffness of phantom device as well as the general lack of steady control made it difficult for users to avoid errors. For example, an operator might want to stop the excavator immediately in case of emergency; however, he might not be able to do this due to the lack of steady control. Also, because the excavator sometimes did not respond well to operator commands, operator may not be able to completely control the excavator at all times to prevent errors from occurring. Further, introducing start and stop points (limit points) on stylus rotation will help reduce operator frustration and improve performance on tasks. For example, when bucket is fully open, it should correspond to the limit of rotation of the stylus in one direction, likewise, when it is fully closed, it should correspond to the limit of rotation of the stylus in the other direction. This way, a point in rotation will be reached when operator knows the bucket is fully opened/extended or when bucket is fully closed/retracted. In other words stylus rotation should stop when bucket is fully open or closed (stylus should rotate 180° so it is exactly mimics the bucket).

6.5 Design Improvements to Haptic-Controlled Excavator Interface

Based on the empirical studies, task analysis, user profile and the usability goals set for the haptic-controlled excavator interface, the following design improvements and modifications were suggested for the haptic-controlled excavator.

(1) The results from the empirical study showed no statistically significant difference in task completion time, number of scoops and number of drops for the three levels of force feedback tested. However, the results also showed users preferred medium force feedback to low force feedback and high force feedback conditions in terms of comfort, ease of use and frustration. Further, users rated medium force feedback lower in terms of mental workload, physical workload, frustration and total workload in NASA TLX assessment. In light of the above, medium force feedback is recommended for the haptic-controlled excavator interface. Thus, a force feedback range of 0.2-1.2N is recommended for use on both the actual haptic-controlled excavator prototype currently under construction at Georgia Institute of Technology, as well as on the laboratory version of the simulator in use at North Carolina Agricultural & Technical State University (NCA&T).

(2) The empirical study identified hand-eye coordination as an issue that affects operator performance, especially for novices, as they struggled to coordinate their eye movement between work area and haptic device. To reduce hand-eye coordination struggle, the haptic device must be positioned on an adjustable stand within the center of operator reach zone, and away from the cabin walls of the excavator as shown in Figure 6.6. As shown in Figure 6.6, the haptic device needs

to be placed within about 30cm from operator elbow, and be adjustable so it could be fit different operators. This recommendation applies to the actual haptic-controlled excavator as well as the simulator version.

(3) The questionnaire from empirical study as well as the user profile results identified shoulder, elbow and wrist fatigue as common problem associated with operation of the haptic device for long period of time. The fatigue associated with operating the haptic device for long hours can be minimized by providing adequate, well placed and comfortable arm rest as shown in Figure 6.7 to provide the needed support. This recommendation applies to both actual haptic-controlled excavator and the simulator version at NCA&T.

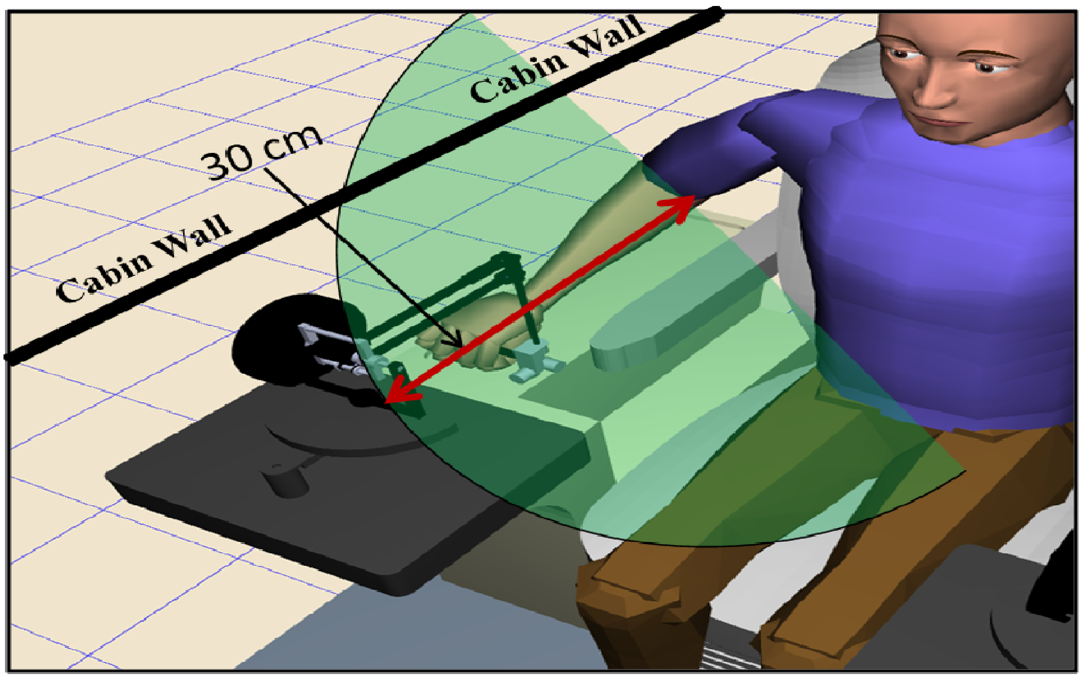

Figure 6.6: Proposed position of haptic device in excavator cabin

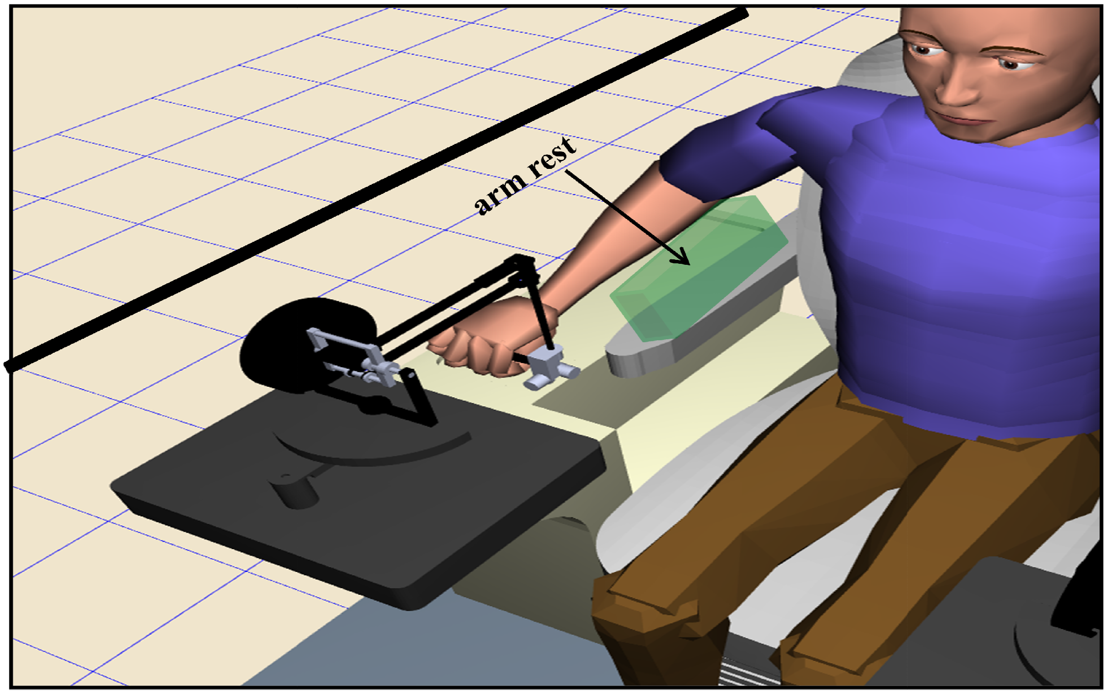

Figure 6.7: Proposed operator arm rest location in excavator cabin design

(4) To provide operator with complete and unobstructed view of the work area, real time video information of the trench collected with bucket mount cameras is proposed to provide a secondary view of work area. The real time video information will be displayed on a screen mounted in the upper right-hand corner of cabin as shown in Figure 6.8 below. This recommendation is proposed for the actual haptic-controlled excavator interface.

(5) In order to provide a realistic and useful feedback to operators, it is recommended that the proportional weight of soil content in bucket be appropriately incorporated into the haptic device. This will provide a natural and intuitive feedback to operators in performing excavation tasks, especially in cases where the operator is unable to see directly from the work area. This will help address the concern of operators' inability to see work zone especially at the bottom of the

trench. Further, this will eliminate the need for practice of excavator operators to depend on human assistants/lookouts commonly encountered in excavation tasks. This recommended for use on both the actual haptic-control excavator interface and the simulator version being used at NCA&T.

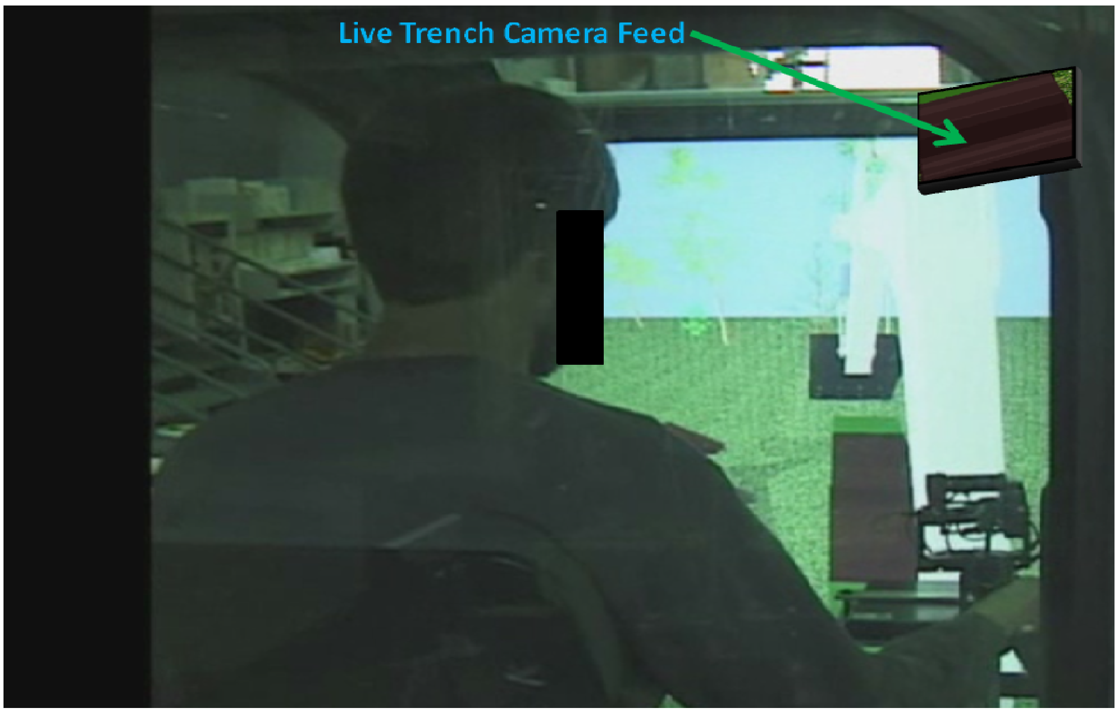

Figure 6.8: Real-time trench information displayed in excavator cabin

(6) In order to provide a realistic and useful feedback to operators, it is recommended that the proportional weight of soil content in bucket be appropriately incorporated into the haptic device. This will provide a natural and intuitive feedback to operators in performing excavation tasks, especially in cases where the operator is unable to see directly from the work area. This will help address the concern of operators' inability to see work zone especially at the bottom of the trench. Further, this will eliminate the need for practice of excavator operators to

depend on human assistants/lookouts commonly encountered in excavation tasks. This recommended for use on both the actual haptic-control excavator interface and the simulator version being used at NCA&T.

(7) Further, to simulate natural noises made by the bucket as it scoops dirt, a scooping sound is recommended to be incorporated in the haptic-controlled excavator. This will provide useful feedback to operators in performing excavation tasks. This will be especially helpful when conducting experiments with the simulated haptic-controlled excavator in laboratory settings, where most of the natural noises in a real environment might be absent.

6.6 Chapter Summary

In this chapter, steps in user-centered design were used to identify changes necessary to improve the haptic controlled excavator interface. Steps outlined include the development of user profile through interviews with excavator operators, and a hierarchical task analysis to identify critical tasks in excavator operation. Results from user profile showed excavator operators generally tend to older, male with lots of on the job experience and novice to moderate computer skills. Task analysis identified moving, positioning, opening, scooping, unloading as the critical tasks performed in excavation task. Next, the results from user profile and task analysis were used in combination with results from empirical studies discussed in Chapters 3 and 4, to develop a set of qualitative and quantitative usability goals for the haptic-controlled excavator. The usability goals were derived to help the design team to focus on “the voice of the user” throughout the design stage, and to identify critical tasks and improve design shortcomings in order to increase effectiveness, efficiency and ease of use of the haptic-controlled excavator.

Usability testing of the haptic-controlled excavator was conducted to identify potential usability problems users might encounter. Usability problems identified include difficulty in controlling bucket due to high sensitivity of stylus, highly unstable boom and bucket, poor synchronization between bucket and stylus, poor bucket response to stylus command, lack of proper arm rest, as well as restricted workspace among others.

Several design improvements and modifications were recommended to increase efficiency, effectiveness and to make the interface more intuitive and user friendly. One key recommendation is the use of medium force feedback in the haptic-controlled excavator interface, since it produced the best operator performance in terms of task completion time and number of scoops required to fill a bin. Further, medium force feedback was rated the least in terms of mental workload, physical workload, frustration and total workload in a NASA TLX assessment by operators. Also, it was rated the most easy to use and least frustrating by operators in a subjective assessment questionnaire. However, it must be noted that the higher operator performance recorded for medium force feedback was statistically not significant.

Other modifications recommended include positioning the haptic device within the center of operator reach zone and away from the cabin walls to reduce struggles in hand-eye coordination and to ensure unobstructed operator control. Further, the provision of appropriate arm rest is recommended to reduce elbow and shoulder fatigue. A boom/bucket mounted camera to provide secondary/redundant cues from work environment is recommended, likewise, the incorporation of bucket weight to provide realistic operator feedback.

CHAPTER 7

SUMMARY, DISCUSSION AND CONCLUSIONS

7.1 Summary

The purpose of this dissertation was to develop a multimodal human-machine interface for haptic-controlled excavator interface. In order to accomplish this goal, this dissertation is organized into six major chapters. Five research questions were addressed in two empirical investigations described in chapters 3 and 4. A summary of each chapter is presented below.

Chapter 1 presented the context, motivation and objective for the dissertation, and defined the problem addressed in this dissertation. The context of the dissertation is multimodal human-machine interfaces, specifically, the use of touch/haptics in the design of haptic-controlled excavator interface. Designing haptic-controlled excavator interface is offered as an alternative to the traditional joystick, lever or pedal human-excavator interfaces currently in use. This has become necessary due to the rapidly aging and shrinking male boomer population from which most traditional excavator operators belong, and the need to attract younger and more diverse excavator operator population. The testbed for this dissertation is the excavator which belongs to the fluid power systems family. Specifically, the goals of the dissertation were:

1. To provide a rationale for using haptic-controlled excavator interface as an alternative to the traditional joystick, levers and pedal human-excavator interface

2. Investigate interference in haptic-controlled excavator interface using empirical study
3. Identify the range of force feedback values that results in best operator performance through empirical evaluation
4. Develop a quantitative excavator-operator model
5. Use user-centered design approach to make design improvements and modifications to the existing haptic-controlled excavator interface.

Chapter 2 presented a literature review of multimodal human-machine interfaces together with their basic theories. The five human senses visual, auditory, haptic, smell and taste were reviewed. Their applications in multimodal human-machine interface design as well as their strength and limitations were also reviewed. The visual, auditory and haptic senses are used in a wide range of applications; however, the senses of taste and smell have limited applications in design mainly due to the fact that the technology is not well developed to allow their use in design. Theories of multimodal human-machine interfaces were reviewed and the rationale for using multimodal haptic-controlled excavator interface (i.e. an interface that uses visual, auditory and haptic senses) as an alternative to the traditional excavator interface (i.e. an interface that uses visual and auditory senses) was outlined.

A strength/advantage in using multimodal human-machine interface to designing human-excavator interface is that

1. It allows information exchange between human and excavator through multiple channels: visual, auditory and haptic;

2. It provides simultaneous exchange of information between human and machine through force feedback.
3. It allows for more immersed interaction between human and machine as the force feedback provides a sense of 'feel' resulting in a more satisfactory interaction.
4. It reduces information overload that is otherwise placed on the visual sense, and
5. It provides redundant cues to the operator.

These attributes make the haptic-controlled excavator interface more efficient, effective and intuitive compared to the traditional joystick, lever and pedal excavator interface.

Chapter 3 presented theories of interference associated with multimodal human-machine interface, as well as challenges and limitations. Multiple resource theory Wicken's (1984) was discussed. An empirical investigation was conducted to assess (i) whether conflict exists between visual, haptic and auditory modalities in the haptic-controlled excavator interface (ii) the impact of interference on operator performance while using the haptic-controlled excavator interface, and (iii) whether operators struggle to coordinate their hand-eye movement. A brief discussion of the research questions are provided in the next section.

Chapter 4 presented a general overview of haptics and its application in fluid power systems as well as its strengths and limitations in multimodal human-machine interface. An empirical study was conducted to investigate (i) whether different levels of force feedback affect operator performance, and (ii) identify the force feedback range that produce best operator performance. A brief discussion of the research questions is provided in the next section.

Chapter 5 presented a quantitative human-excavator model using control theory approach of modeling human performance in pursuit or tracking tasks. Several control theory models used in characterizing operator performance in manual tracking tasks were discussed. The basic concept of the control theory is its ability to compare the control behavior of humans to that of inanimate automatic feedback control systems in order to model operator behavior. This is achieved by developing a mathematical analysis of the dynamic systems of the human operator and the mechanisms for achieving a desired steady state under changing internal and external conditions.

A conceptual human-excavator model together with a control model of both the traditional excavator and the haptic-controlled excavator interface were developed. The models were implemented in *MatLab,* and their stabilities were compared using their Bode and Nyquist plots. Analysis of the models showed that both the traditional excavator and the haptic-controlled excavator interfaces were stable. However, the haptic-controlled excavator interface was found to be more stable than the traditional excavator interface, providing further proof for the rationale to use the haptic-controlled excavator interface as an alternative to the traditional excavator interface.

Chapter 6 presented the user-design centered approach used to identify design problems in the haptic-controlled excavator interface as well as the necessary modifications needed to improve the interface. The user-centered design approach ensured that 'the voice of the user' was at the center of the design throughout the design process. User profile developed through interviews showed that excavator operators are old, male, experienced with long work hours. Hierarchical task analysis was used to identify critical task performed in excavation. Qualitative and quantitative usability goals for the haptic-controlled excavator interface were developed, and usability evaluation was conducted. Design modifications and changes to improve the interface were proposed.

7.2 Summary of Empirical Studies

Two empirical studies were conducted in this dissertation to investigate interference in multimodal human-machine interface and the impact of force feedback on operator performance. The summary of results of the two empirical studies that address each of the five research questions in Chapters 3 and 4 are discussed below.

7.2.1 Question 1. Conflict Detection.

This question was answered by analyzing eye tracking data (fixation count and length) for tasks that depend on auditory, visual and haptic cues for expert and novice operators. The results showed that both mean number of fixation count and mean fixation lengths were higher for novice operators than they were for expert operators. However, mean number of fixation count outside the area of interest (AOI), and the length of fixation outside AOI were higher for novice operators than for expert operators. Further, novices were nearly twice as likely (3.78 vs. 2.0) to look outside the area of interest while performing the task compared to experts. The results showed that, while there was no statistically significant difference in fixation count outside AOI between experts and novices, *(z=48.0*, and *p=0.9054)*, within the area of interest, there was a statistically significant difference in fixation count in the area on interest between experts and novices *(z=104.50*, and *p=0.018)*. Similarly, the results showed that in the area of interest, there was a statistically significant difference in fixation length between experts and novices, *(z=23.00*, and *p=0.0398)*, whereas outside the area of interest, there was no statistically significant difference in fixation length between experts and novices *(z =165.00*, and *p=0.3988)*. The significant difference in the higher number of fixation count and fixation length within the area of interest by experts may be due to the fact that novice operators had harder time keeping their eyes focused in the task area compared to expert operators.

7.2.2 Question 2. Impact of Conflict on Performance.

This question was answered by comparing the performance of experts and novices using task completion time, number of scoops required to fill a bin, and number of scoops dropped outside of bin. Results from the empirical study showed that there was a statistically significant difference between experts and novices in task completion time *(z=18,* and *p=0.0047)*, number of scoops required to fill a bin *(z=21,* and *p=0.0240)*, however, there was no significant difference in the number of drops outside of the bin *(z=37.50, p=06059)*. This may be due to the fact that experts had a higher fixation count in the area of interest than novices, which may be attributed to the fact that experts were able to focus their attention in the work area where the actual excavation task took place, while novices wandered in and out of the area of interest. Further, the results also show that, training can be used to greatly improve the performance of novice operators.

7.2.3 Question 3. Hand-Eye Coordination.

This question was answered by analyzing eye-tracking data on and outside the area of interest together with scan paths data. Results showed that within the area of interest, there was a statistically significant difference between experts and novices in fixation count *($z=104.50$, and $p=0.0118$),* and fixation length ($z=23.0$, and $p=0.0398$). However, outside the area of interest, there was no statistically significant difference between experts and novices in fixation count *($z=48.0$,* and *$p=0.9054$)*, and fixation length *($z=16.5$*, and *$p=0.3988$)*. Further, gaze plots show that novice operators had more difficulty coordinating their hand-eye movement than did expert operators.

7.2.4 Question 4. Impact of Force Feedback on Performance.

This question was answered by comparing the performance of operators under different force feedback conditions. The results showed a statistically significant difference in task completion time between the different levels of force feedback *($H=9.94207$, 2 d.f,* and *$p=0.0242$)*. Also, a statistically significant difference between the different force feedback conditions *($H=9.031$, 2 d.f,* and *$p=0.0288$)* was obtained in terms of number of scoops. However, there was no statistically significant difference between the different force feedback conditions *($H=1.6549$, 2 d.f,* and *$p=0.6470$)* in terms of the number of drops outside the bin. Thus, the results showed that, the levels of force feedback had a significant effect on operator performance.

7.2.5 Question 5. Levels of Force Feedback and Performance.

This question was answered by comparing the performance of operators under low, medium and high force feedback conditions. The results showed no statistically significant difference in task completion time *(H=0.9834, 2 d.f,* and *p=0.6116)* between the means of low force feedback, medium force feedback and high force feedback. Likewise, there was no statistically significant difference in the mean number of scoops required to fill up a bin *(H=0.2093, 2 d.f, and p=0.9006)* for three force feedback conditions. However, in a subjective questionnaire completed after the test, participants rated medium force feedback as the least frustrating and most easy to use. Similar results were also obtained from NASA TLX workload assessment in which participants rated medium force feedback highest in terms of mental workload, physical workload, frustration and total workload.

7.3 General Discussion

The haptic-controlled excavator interface is a multimodal human-machine interface that is being developed as an alternative to the traditional human-excavator interface. It exploits the benefits of visual, auditory and haptic modalities in a user interface to synergistically compliment each others' weakness in order to yield a more effective and intuitive interface. In particular, the simultaneous exchange of information between the human and machine is exploited to produce a more immersed interaction. Results from the empirical study showed that using force feedback indeed improved operator performance significantly for task completion time *(H=9.94207, 2 d.f,* and *p=0.0242)* and number of scoops to fill a bin *(H=9.031, 2 d.f,* and *p=0.0288)*. However, on the question of which level of force feedback yields best operator performance, the results did not show a significant difference, though the author recommends medium force feedback range since both NASA TLX assessment and subjective questionnaire showed operators preferred medium force feedback.

In spite of the potential benefits associated with multimodal human-excavator interface discussed, results from the empirical study showed conflicts do exist between the visual, auditory and haptic modalities in a haptic-controlled excavator interface. The impact of the conflict on performance is greater on novices than experts. The impact of this can be seen in the poor hand-eye coordination observed in novices. The author has provided several design modifications to address these issues in Section 6.5.

To provide a rationale for the use of haptic-controlled excavator interface as an alternative to the traditional human-excavator interface, quantitative models of both haptic-controlled excavator and traditional human-excavator interfaces were developed using control theory to model the human operator as a control system. The models were implemented in Matlab, and while both models showed stability, the haptic-controlled excavator model was found to be more stable than the traditional human-excavator model. Also, models of visual only, auditory only and haptic only sensory cue processing were developed and compared. The results showed that visual only information processing had the highest gain and phase angle shift followed by haptic only and auditory only cue processing respectively.

7.4 Contributions

This dissertation contributes towards the enhancement of multimodal human-machine interface theory, specifically the development of haptic-controlled excavator interface as an alternative to traditional human-excavator interface. The following are key contributions.

1. The dissertation used empirical investigation to assess conflicts in haptic-controlled excavator interface and provided design modifications necessary to mitigate the impact of potential conflicts on operator performance, and identified the level of force feedback in the haptic-controlled excavator interface that yields best operator performance.
2. The dissertation developed a conceptual framework for the interaction between human operator and excavator which provided the basis for using multiple sensing modalities in the haptic-controlled excavator interface.
3. The dissertation developed a quantitative model to aid in predicting operator performance in when using the haptic-controlled excavator and traditional excavator interfaces. Results from the quantitative model help provide a rationale for using the haptic-controlled excavator as an alternative to the traditional human-excavator interface.
4. The dissertation developed a user-centered design approach that brought the user to the center of the haptic-controlled excavator interface design effort, and serves as template for developing haptic-controlled interfaces for other fluid power system in the future.

7.5 Limitations

As is often the case with most research, this study had several inherent limitations. First, the two empirical investigations were conducted in a laboratory environment using a simulated haptic-controlled excavator and a computer screen as the task environment. This set up lacked the attributes of real excavator environment such as vibration and engine noise. It will, therefore, be useful to replicate the experiment in a real excavator cabin with characteristics similar to those found in excavation task environments.

Both empirical studies recruited students as participants in conducting the investigation. This participant pool is not representative of the general excavator operator population which is usually older and male. Also, due to time and monetary constraints, only small number of participants 20 novices and 4 experts were recruited for the first empirical study; likewise, only 20 participants were recruited for the second empirical study. Since the haptic-controlled excavator interface is a new design, it was difficult to determine who an expert was. In this conducting this study, a person was considered an expert if he/she was a member of the CCEFP team, knowledgeable about the haptic-controlled excavator, had experience with and used it several times in the past.

Also, since the experiments were conducted in a simulated laboratory environment, the impact and/or contribution of sound/noise to operator performance was difficult to assess. In this study, the only use of sound was to provide an auditory alert to participants when excavation task was complete. It will be useful to replicate the experiment in an environment where quantifiable and measurable auditory cues could be used to assess its impact on operator performance during excavation task.

In addition, using control theory to develop a quantitative model of the interaction between the operator and the haptic-controlled human excavator, several assumptions were made about the conceptual model of the operator-excavator interaction. For example, the human operator was modeled as a purely mechanical system in order to develop the quantitative model. This may impact model adequacy and performance as the human is not purely mechanical as assumed in the model. Another limitation of this study is the fact that author was unable to validate the quantitative model due to time and resource constraints.

7.6 Recommended Future Research

The following have been proposed as logical follow up studies to the current study in a number of directions with meaningful implications.

1. First, to make the results from the empirical investigations more realistic and to ensure that the results could be applied in real life environments; there is the need to replicate the experiments using haptic-controlled excavator with similar layout and attributes found in real excavation task environments.
2. The participant pool for the current study was drawn from the student population which did not reflect the user population of excavator operators. Therefore, there is the need to replicate the study using a larger number of participants drawn from user group that resembles the true excavator operator population and investigate whether this has an impact on operator performance.
3. The current study made little use of auditory feedback as the laboratory simulated haptic-controlled excavator environment made it difficult to practically incorporate quantifiable auditory information into the experimental

design. It will be useful to conduct the experiment using quantifiable auditory feedback in order to assess the impact of auditory feedback on operator performance during excavation task, and whether conflict might arise due to the use of auditory feedback.

4. The current study is focused on implementing the haptic-controlled interface on the excavator; however, since CCEFP is involved in multiple test bed projects, it will be useful to implement the haptic-controlled interface on the other test beds currently being developed by the CCEFP group.
5. Another logical follow up study to the quantitative model developed in the current study is to validate the model by conducting simulated experiments using the developed model and comparing the results with empirical results obtained from user experiments. This will establish the validity and reliability of the quantitative model developed for the human-excavator interaction in a haptic-controlled excavator interface.

REFERENCES

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. In Bainbridge, W (editor). *Encyclopedia of Human-Computer Interaction*. Thousand Oaks: Sage Publications.

Alleyne, A., & Liu, R. (2000). A simplified approach to force control for electro-hydraulic systems. *Control Engineering Practice, 8*(12), 1347-1356.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review, 89*(4), 369.

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences, 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *The psychology of learning and motivation, 8*, 47-89.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of personality and social psychology, 100*(3), 407.

Benel, D. C. R., Ottens, D., & Horst, R. (1991). Use of an eyetracking system in the usability laboratory. *Proceedings of the Human Factors Society 35th Annual Meeting*. Santa Monica, CA. 461-465.

Bernsen, N. O. (1994). Foundations of multimodal representations: a taxonomy of representational modalities. *Interacting with computers, 6*(4), 347-371.

Bernsen, N. O. (1997). Defining a taxonomy of output modalities from an HCI perspective. *Computer standards & interfaces, 18*(6-7), 537-553.

Biggs, S. J., & Srinivasan, M. A. (2002). Haptic interfaces. *Handbook of virtual environments*: *Design, Implementation, and Applications*. K.M. Stanney, ed., Lawrence Erlbaum, NJ. 93-116.

Binsted, G., Chua, R., Helsen, W., & Elliott, D. (2001). Eye-hand coordination in goal-directed aiming. *Human Movement Science, 20*(4-5), 563-585.

Bjelland, H., & Tangeland, K. (2007). User-centered design proposals for prototyping haptic user interfaces. *Haptic and Audio Interaction Design*, 110-120.

Bjur, J. J. (1998). Auditory Icons in an Information Space. *Department of Industrial Design, School of Design and Craft, Goteberg University, Sweden.*

Bowen, K., & Marcia, K. (2006). Adaptation of haptic interfaces for a labview-based system dynamics course. In *HAPTICS*, IEEE, 147-152.

Brave, S., & Dahley, A. (1997). inTouch: a medium for haptic interpersonal communication. *Ext. Abstracts CHI 1997, ACM Press*, 363-364

Brewster, S. A. (1997). Using non-speech sound to overcome information overload. *Displays, 17*(3-4), 179-189.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual perception: Physiology, psychology, & ecology*: Hove, UK. Psychology Press.

Buck, L., & Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell, 65*(1), 175-187.

Burdea, G., & Brooks, F. P. (1996). *Force and touch feedback for virtual reality*: Wiley New York.

Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM, 23*(7), 396-410.

Carter, R. (2008). Hydraulic excavator overhauls: When and what to plan for. *Engineering & Mining Journal, 113*(5), 50-52.

Chang, D., Dooley, L., & Tuovinen, J. E. (2002). *Gestalt theory in visual screen design: a new look at an old subject*. Proceedings of the 7th World Conference on Computers in Education, Demark. Vol.8, 5-12.

Cooper, B. (1998). Driving on the surface of Mars using the Rover Control Workstation. In *Proceedings of SpaceOps 1998*, Tokyo, Japan.

Deng, F. (1999). *Control Models of the Human Operator in multi-sensoty information Processing Task Environment.* Unpublished Masters' thesis, North Carolina A&T State Univerisrty, Greensboro, NC.

Deng, F., & Ntuen, C. (1998). *A study of adaptive control of the human operation in a computer human-machine system*. Paper presented at the Industrial Engineering Theory and Application Conference, Hong Kong.

Dumas, B., Lalanne, D., & Oviatt, S. (2009). Multimodal interfaces: A survey of principles, models and frameworks. D. Lalanne and J. Kohlas (Eds.): *Human Machine Interaction*, LNCS 5440. Verlag, Berlin. pp 3-26.

Elkind, J. I. (1956). *Characteristics of simple manual control systems*. M.I.T. Lincoln Lab. Lexington, Mass., Tech. Rept. No. 111

Erickson, R. P. (1982). Studies on the perception of taste: do primaries exist? *Physiology & Behavior, 28*(1), 57-62.

Fang, J. (1997). *Performance models of the human operator in discrete and conclusion control tasks.* Masters, North Carolina A&T State University, Greensboro, NC.

Frankel, J. G. (2004). *Development of a haptic backhoe testbed.* MS thesis, The Georgia Institute of Technology, G.W. Woodruff School of Mechanical Engineering

Fultz, J. (1999). Theory of Gestalt Psychology. Retrieved on 12/07/2010 from http://users.anderson.edu/~jfultz/gestalt.html website

Garcia-Ruiz, M. A., El-Seoud, S. A., Edwards, A., Al-Jaam, M. J., & Aquino-Santos, R. (2008). *Integrating the Sense of Smell in an Educational Human-Computer Interface*. Paper presented at the *Interactive Computer Aided Learning*, Villach, Austria.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction, 4*(1), 67-94.

George, G. R. (2009). *New methods of mathematical modeling of human behavior in the manual tracking task.* PhD Dissertation, State University of New York at Binghamtom.

Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. *The mind's eye: Cognitive and applied aspects of eye movements research*, 493-516.

Graziano, M. S. A., & Gross, C. G. (1993). A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Experimental Brain Research, 97*(1), 96-109.

Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., & Cajander, Å. (2003). Key principles for user-centred systems design. *Behaviour and Information Technology, 22*(6), 397-409.

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE, 85*(1), 6-23.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research, in P. A. Handcock and N. Meshkati (eds), Human Mental Workload. New York, Elsevier, 5-39.

Hayward, V., Astley, O. R., Cruz-Hernandez, M., Grant, D., & Robles-De-La-Torre, G. (2004). Haptic interfaces and devices. *Sensor Review, 24*(1), 16-29.

Hayward, V., & MacLean, K. E. (2007). Do it yourself haptics: Part I. *Robotics & Automation Magazine, IEEE, 14*(4), 88-104.

Hess, R. A. (1965). The human operator as an element in a control system with time varying dynamics: AFFDL-FDCC-TM, 65-34.

Hess, R. A. (1985). A model-based theory for analyzing human control behaviour. *Advances in Man-Machine Systems Research, 2*, 129-175.

Hess, R. A. (1995). Modeling the Effects of Display Quality Upon Human Pilot Dynamics and Perceived Vehicle Handling Qualities. *IEEE Transactions on Systems Man and Cybernetics, 25*(2), 338-344.

Hess, R. A., & Shipman, V. C. (1965). Early experience and the socialization of cognitive modes in children. *Child development*, 869-886.

Hone, G., & Stanton, N. (2004). HTA: The development and use of tools for hierarchical task analysis in the Armed Forces and elsewhere. *Human Factors Integration Defense Technology Centre.*

Horrey, W. J., & Wickens, C. D. (2003). *Multiple resource modeling of task interference in vehicle control, hazard awareness and in-vehicle task performance*. Proceedings of the 2nd International Driving Symposium on Huamn Factors in Driver Assessment Training and Vehicle Design, Park City, UT, *pp* 7-12.

Hosman, R., & Stassen, H. (1999). Pilot's perception in the control of aircraft motions. *Control Engineering Practice, 7*(11), 1421-1428.

Howard, I. P. (1996). Alhazen's neglected discoveries of visual phenomena. *Perception. London, UK. 25*, 1203-1218.

Hurtig, T., & Jokinen, K. (2006). Modality fusion in a route navigation system. In *Proceedings of Workshop on Effective Multimodal Dialogue Interfaces*, Sidney, Australia.

Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, Amsterdam *2*(3), 573-605.

Jacobson, R. D., Kitchin, R., & Golledge, R. (2002). 25 Multi-modal virtual reality for presenting geographic information. *Virtual reality in geography, 1*, 382.

Jaimes, A., & Sebe, N. (2007). Multimodal human-computer interaction: A survey. *Computer Vision and Image Understanding, 108*(1-2), 116-134.

Johnson, K. O., & Phillips, J. R. (1981). Tactile spatial resolution. I. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology, 46*(6), 1177.

Johnson, M. J., Van der Loos, H. F. M., Burgar, C. G., Shor, P., & Leifer, L. J. (2005). Experimental results using force-feedback cueing in robot-assisted stroke therapy. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 13*(3), 335-348.

Jokela, T., Iivari, N., Matero, J., & Karukka, M. (2003). *The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11*.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology, 8*(4), 441-480.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kantowitz, B. H., & Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships*: Wiley, New York.

Katz, D., & Krueger, L. E. (1989). *The world of touch*: Edited by Krueger, LE. Hillsdale, New Jersey. Erlbaum Associates, Inc.

Kaye, J. J. (2004). Making Scents: aromatic output for HCI. *interactions, 11*(1), 48-61.

Kaye, J. N. (2001). *Symbolic olfactory displays.* Masters thesis, Media Lab. Massachusetts Institute of Technology, Cambridge, MA.

Keehner, M. (2008). Conflicting Cues from Vision and Touch Can Impair Spatial Task Performance: Speculations on the Role of Spatial Ability in Reconciling Frames of Reference. *Spatial Cognition VI. Learning, Reasoning, and Talking about Space*, 188-201.

Kirwan, B. E., & Ainsworth, L. K. (1992). *A guide to task analysis*: Taylor & Francis, Basingstoke, UK.

Kleinman, D., Baron, S., & Levison, W. (1970). An optimal control model of human response part I: Theory and validation. *Automatica, 6*(3), 357-369.

Kleinman, D., Baron, S., & Levison, W. (1971). A control theoretic approach to manned-vehicle systems analysis. *Automatic Control, IEEE Transactions on, 16*(6), 824-832.

Koffka, K. (1955). *Principles of Gestalt psychology* (Vol. 7): Psychology Press.

Koka, K. (1935). Principles of Gestalt psychology. *Harcourt, Brace & World, New York.*

Kontz, M., & Book, W. (2007). Electronic control of pump pressure for a small haptic backhoe. *International Journal of Fluid Power, 8*(2), 5.

Kramer, G. (1994). *An introduction to auditory display*. In Kramer G (ed) Auditory display, sonification, audification, and auditory interfaces, Addison-Wesley, Reading, Massachusetts, pp 1-78

Krapichler, C., Haubner, M., Lösch, A., Schuhmann, D., Seemann, M., & Englmeier, K. H. (1999). Physicians in virtual environments multimodal human-computer interaction. *Interacting with Computers, 11*(4), 427-452.

Legin, A., Rudnitskaya, A., Seleznev, B., & Vlasov, Y. (2005). Electronic tongue for quality assessment of ethanol, vodka and eau-de-vie. *Analytica Chimica Acta, 534*(1), 129-135.

Leinonen, L., Hyvärinen, J., Nyman, G., & Linnankoski, I. (1979). I. Functional properties of neurons in lateral part of associative area 7 in awake monkeys. *Experimental Brain Research, 34*(2), 299-320.

Locher, P. J. (1982). Influence of Vision on Haptic Encoding Processes. *Perceptual and Motor Skills, 55*(1), 59-74.

MacLean, K. E., & Hayward, V. (2008). Do it yourself haptics: Part ii [tutorial]. *Robotics & Automation Magazine, IEEE, 15*(1), 104-119.

Massaro, D. W. (1987). Speech perception by ear and eye: A paradigm for psychological inquiry: *Hillsdale, New Jersey. Lawrence Erlbaum.*

Massaro, D. W. (2004). A framework for evaluating multimodal integration by humans and a role for embodied conversational agents. In *Proceedings of International Conference on Multimodal Interfaces.* State College, PA. pp 24-31.

McLaughlin, M. L., Hespanha, J. P., & Sukhatme, G. S. (2002). *Touch in virtual environments*: Prentice Hall, New Jersey.

McRuer, D. (1980). Human dynamics in man-machine systems. *Automatica, 16*(3), 237-253.

McRuer, D. T. & Krendel, E.S (1974). Mathematical models of human pilot behavior. Advisory Group on Aerospace Research and Development. AGARDograph 188

McRuer, D. T., & Jex, H. R. (1967). A review of quasi-linear pilot models. *Human Factors in Electronics, IEEE Transactions on*(3), 231-249.

McRuer, D. T., & Krendel, E. S. (1959). The human operator as a servo system element 1. *Journal of the Franklin Institute, 267*(5), 381-403.

Minogue, J., & Jones, M. G. (2006). Haptics in education: exploring an untapped sensory modality. *Review of Educational Research, 76*(3), 317.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological review, 86*(3), 214.

Newell, F. D. (1967). In-Flight Measurement of Human Response Characteristics. AIAA Flight Test Simulation and Support Conference Paper No 67-240.

Norman, D. A. (1988). *The psychology of everyday things*. New York. Basic Books.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive psychology, 7*(1), 44-64.

Norman, D. A., & Draper, S. W. (Eds.). (1986). *User centered system design; new perspectives on human-computer interaction*: Lwarence Erlbaum Associates Inc. Hillsdale. N.J.

Oakley I. (2003). *Haptic Augmentation of the Cursor: Transforming Virtual Actions into Physical Actions.* PhD Dissertation, University of Glasgow, Glasgow

Oakley, I., Brewster, S., & Gray, P. (2003). Haptic augmentation of the cursor: Transforming virtual actions into physical actions. In *Computing Science*, 283.

Osafo-Yeboah, B., Delpish, R., & Jiang, X. (2009). Investigation of Operator Behavior Using Haptic Controlled Backhoe Simulator. *The Journal of Management and Engineering Integration, 1, No. 2, pp 9-14* (2), 9-14.

Osafo-Yeboah, B., Elton, M. D., Jiang, X., & Book, W. J. (2010). Usability Evaluation of a Coordinated Excavator Controller with Haptic Feedback. Proceedings of the IERC Conference

Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the ACM, 42*(11), 74-81.

Oviatt, S. (2006). *Human-centered design meets cognitive load theory: designing interfaces that help people think*. In Porceedings of the 14th Annual International ACM Conference on Multimedia. Santa Barbara, CA, USA. pp. 871–880.

Oviatt, S., Cohen, P., Wu, L., Vergo, J., Duncan, L., Suhm, B., Landay, J. (2000). Designing the user interface for multimodal speech and pen-based gesture applications: state-of-the-art systems and future research directions. *Human-Computer Interaction, 15*(4), 263-322.

Oviatt, S. (2003). Multimodal interfaces. *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*, 286-304.

Parkes, A., & Coleman, N. (1990). Route guidance systems: A comparison of methods of presenting directional information to the driver. *Contemporary ergonomics*, 480-485.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*(2), 220.

Poole, A., & Ball, L. J. (2005). Eye tracking in human-computer interaction and usability research: current status and future prospects. *Encyclopedia of human computer interaction*, 211-219.

Popescu, V., Burdea, G., Bouzit, M., & Girone, M. (1999). PC-based telerehabilitation system with force feedback. *Medicine meets virtual reality: the convergence of physical & informational technologies: options for a new era in healthcare, 62*, 261.

Prasanna, R., Yang, L., & King, M. (2009). GDIA: a Cognitive Task Analysis Protocol to Capture the Information Requirements of Emergency First Responders. *Proceedings of the 6th International ISCRAM Conference. Gothenburg, Sweden.*

Quek, F. K. H. (1995). Eyes in the interface. *Image and Vision Computing, 13*(6), 511-525.

Rayner, K., & Pollatsek, A. (1989). The psychology of reading: *Hillsdale, New Jersey. Lawrence Erlbaum*.

Reiser, O. L. (1936). Principles of Gestalt Psychology (Vol. 45, pp. 412-415): JSTOR.

Révész, G. (1950). Psychology and art of the blind. Longmans Green, New York.

Rouse, W. B. (1980). Systems engineering models of human-machine interaction: *North-Holland, New York.*

Rouse, W. B. (1984). Advances in man-machine systems research (Vol. 1). *Greenwich, CT. Jai Press.*

Rudmann, D. S., McConkie, G. W., & Zheng, X. S. (2003). Eyetracking in cognitive state detection for HCI. In *Proceedings of International Conference on Multimodal Interaction.* Vancouver, Canada.

Salisbury, K. (1995). Haptics: the technology of touch. *HPCwire.* Available: http://www.sensable.com/haptics/haptwhpp.html

Sarter, N. B. (2006). Multimodal information presentation: Design guidance and research challenges. *International journal of industrial ergonomics, 36*(5), 439-445.

Sathian, K., & Zangaladze, A. (2002). Feeling with the mind's eye: contribution of visual cortex to tactile perception. *Behavioural brain research, 135*(1-2), 127-132.

Schiessl, M., Duda, S., Thölke, A., & Fischer, R. (2003). Eye tracking and its application in usability and media research. *MMI-interaktiv Journal, 6*, 41-50.

Schomaker, L., Nijtmans, J., Camurri, A., Lavagetto, F., Morasso, P., Benoît, C., Adjoudani, A. (1995). A taxonomy of multimodal interaction in the human information processing system. *Technical Report*, *Esprit Project 8579, MIAMI.*

Sharma, R., Pavlovic, V. I., & Huang, T. S. (1998). Toward multimodal human-computer interface. *Proceedings of the IEEE, 86*(5), 853-869.

Sheridan, T. B., & Ferrell, W. R. (1974). *Man-machine systems; Information, control, and decision models of human performance*: The MIT Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review, 84*(2), 127.

Shoeib, M., Harner, T., & Vlahos, P. (2006). Perfluorinated chemicals in the Arctic atmosphere. *Environmental science & technology, 40*(24), 7577-7583.

Smith, R. (1976). A Unified Theory for Pilot Opinion Rating. *Twelfth Annual Confprpnce on Manual Control, NASA TM X-73170,. pps*, 542-558.

Smith, R. E., & Buchholz, L. M. (1991). Multiple resource theory and consumer processing of broadcast advertisements: An involvement perspective. *Journal of Advertising*, 1-7.

Srinivasan, M. (1995). What Is Haptics. TouchLab. Massachusetts Institute of Technology. Retrieved from: http://touchlab.mit.edu

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics, 37*(1), 55-79.

Stenfelt, S., Hato, N., & Goode, R. L. (2004). Round window membrane motion with air conduction and bone conduction stimulation. *Hearing research, 198*(1-2), 10-24.

Sternberg, R. J., Mio, J., & Mio, J. S. (2009). *Cognitive psychology*: Wadsworth Pub Co, Belmont, CA. USA.

Stone, R. (2001). Haptic feedback: A brief history from telepresence to virtual reality. *Haptic Human-Computer Interaction*, 1-16.

Strauss, E. (1999). Feeling the Future. *Scientific American Presents*. Vol. 10 No. 3, pp 44-47.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119*(2), 176.

Szalma, J., & Hancock, P. (2002). On mental resources and performance under stress. *US Army Research Office, MURI-OPUS. White Paper*.

Thiran, J. P., Bourlard, H., & Marqués, F. (2009). *Multimodal Signal Processing: Theory and applications for human-computer interaction*: Academic Press.

Truax, B. (2001). Handbook of Acoustic Ecology, (CD-ROM version). *Computer Music Journal, 25*, 93-94.

Turk, M., & Kölsch, M. (2004). Perceptual interfaces. *Emerging Topics in Computer Vision, Prentice Hall, New Jersery*.

Tustin, A. (1944). An investigation of the operator's response in manual control of a power driven gun. *Metro-Vickers Electric, SS Memo, 169*.

Verry, R. (1998). Don't take touch for granted: An interview with Susan Lederman. *Teaching of Psychology, 25*(1), 64-67.

Vitense, H. S., Jacko, J. A., and Emery, V. K. (2003). Multimodal feedback: An assessment of performance and mental workload. Accepted to appear in *Ergonomics*. 46: 1-3, 68-87

Wickens, C. D. (1976). The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance, 2*(1), 1.

Wickens, C. D. (1980). The structure of attentional resources. *Attention and performance VIII, 8*.

Wickens, C. D. (1981). Processing Resources in Attention, Dual Task Performance, and Workload Assessment: *In R. Nickerson, & R. Pew (Eds.), Attention and Performance,* Vol. VIII, (pp. 239-257). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Wickens, C. D., Gordon, S. E., & Liu, Y. (2004). *An introduction to human factors engineering*: Pearson Prentice Hall, New Jersey.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 30*(5), 599-616.

Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 25*(2), 227-248.

Woods, A. T., & Newell, F. N. (2004). Visual, haptic and cross-modal recognition of objects and scenes. *Journal of Physiology-Paris, 98*(1-3), 147-159.

APPENDIX

Appendix A

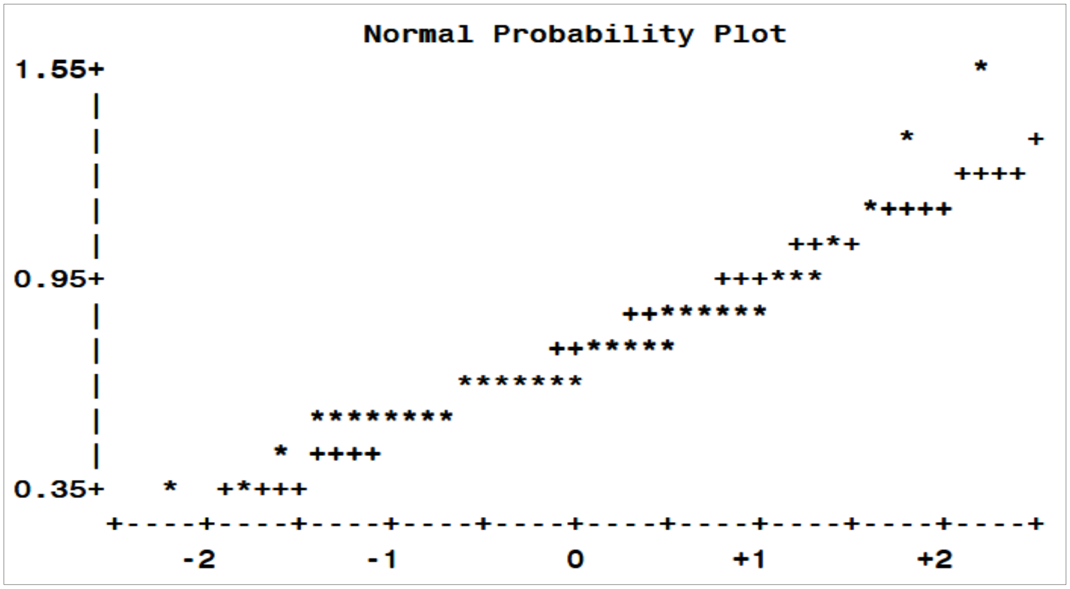

Figure A.1: Normal probability plot for fixation length

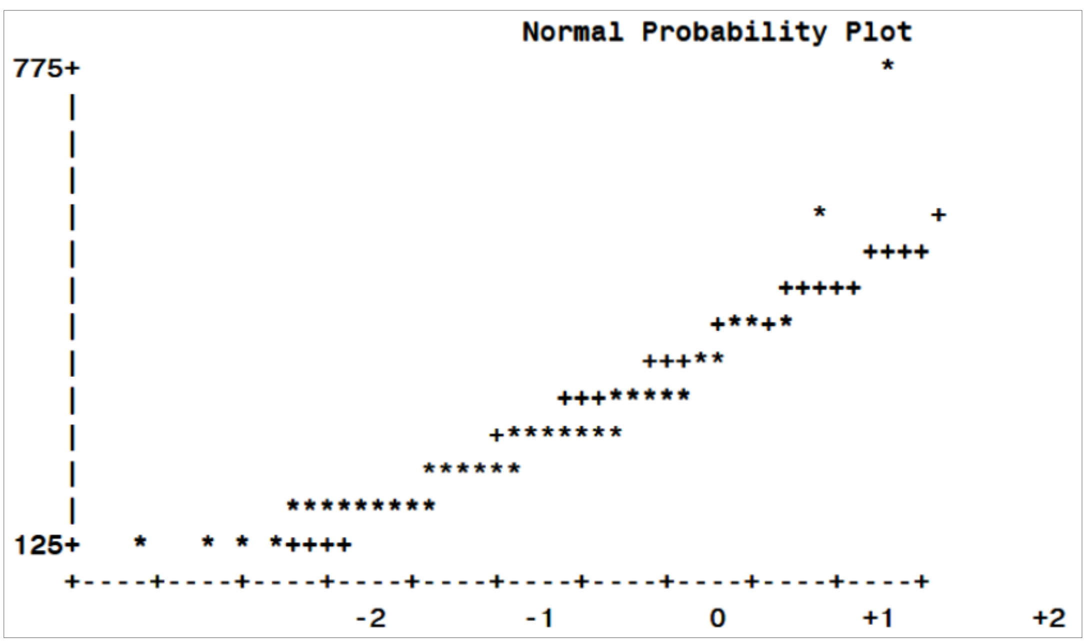

Figure A.2: Normal probability plot for fixation count

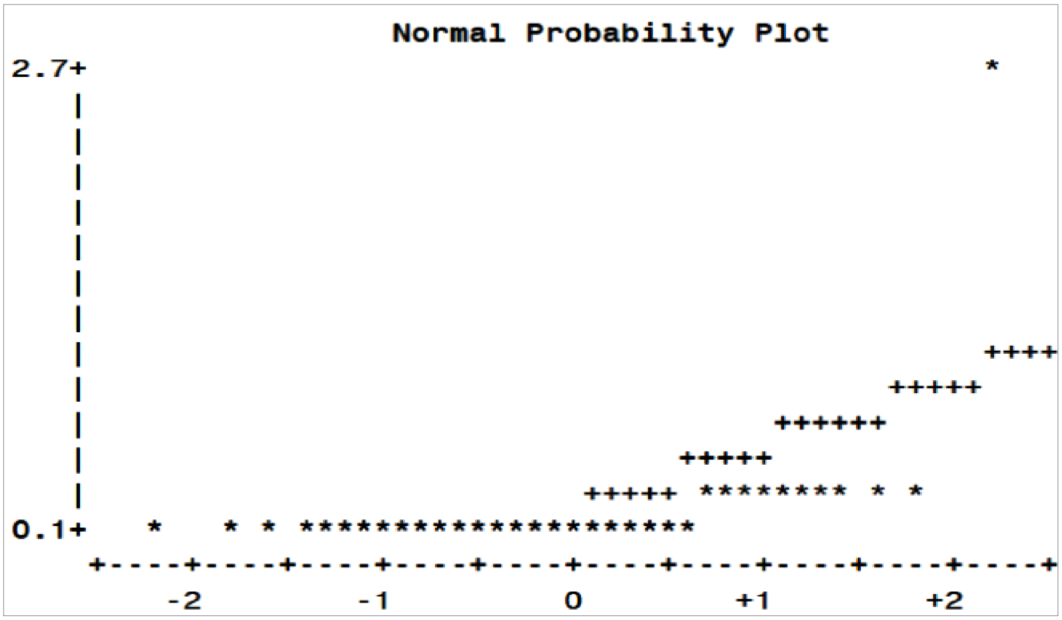

Figure A.3: Normal probability plot for fixation duration

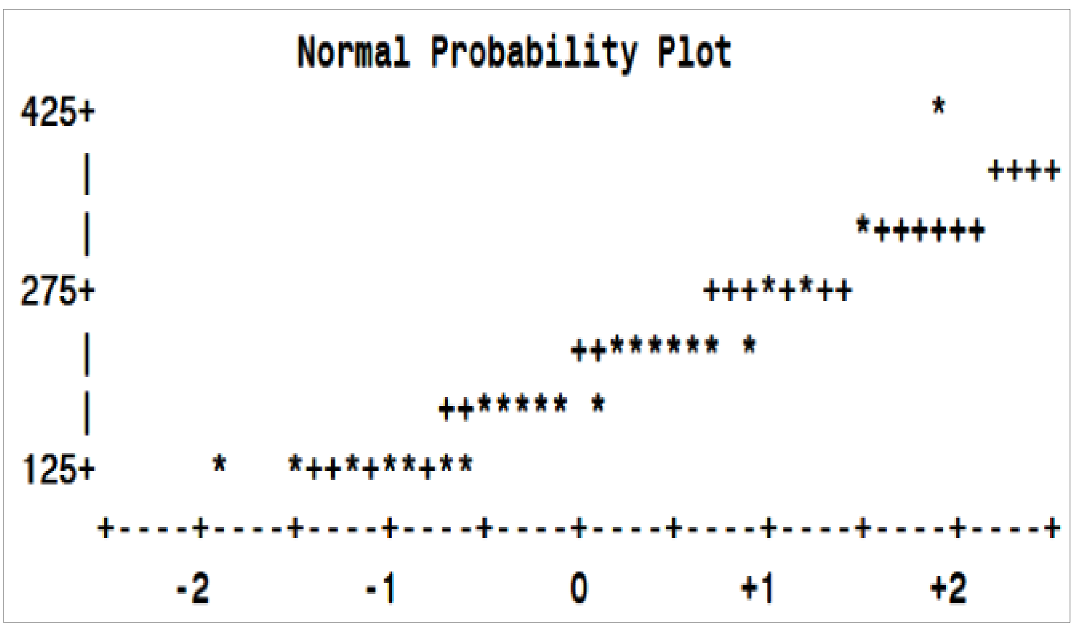

Figure A.4: Normal probability plot for task completion time

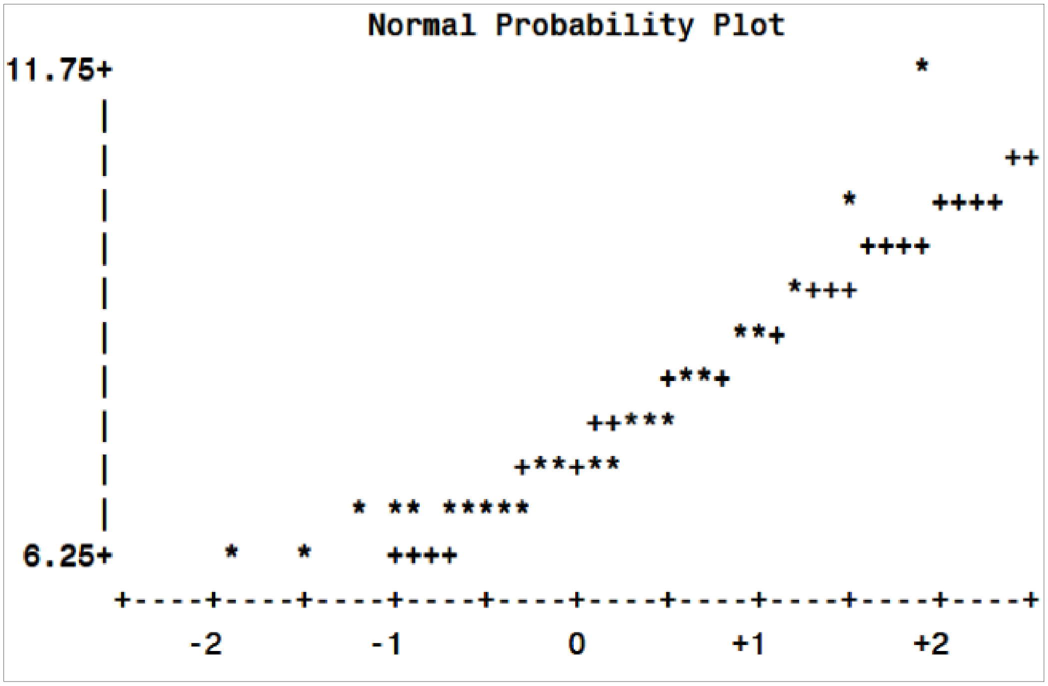

Figure A.5: Normal probability plot for mean # of scoops

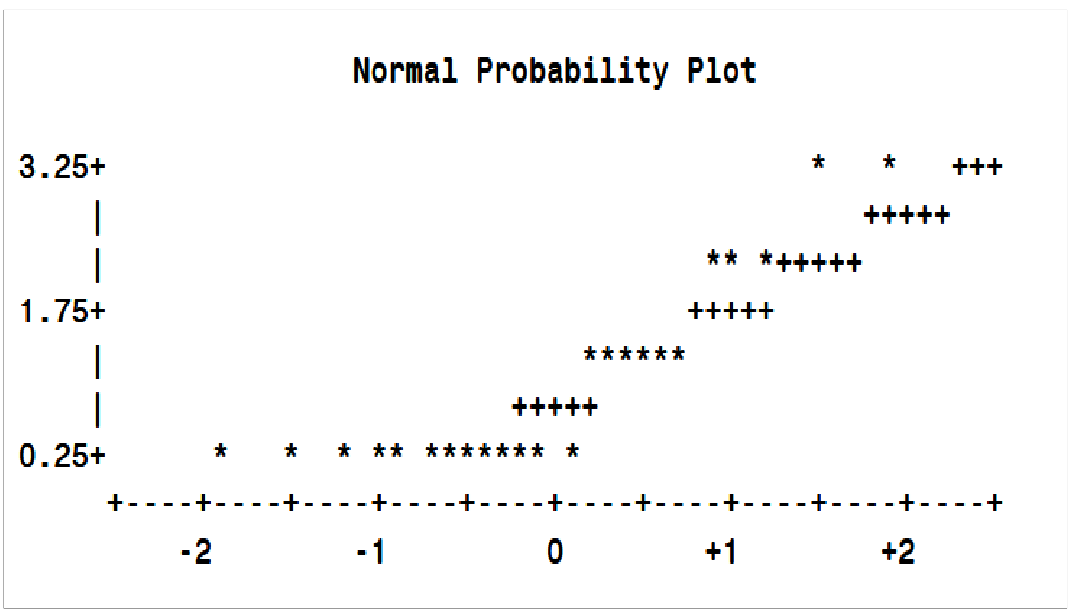

Figure A.6: Normal probability plot for mean # of drops

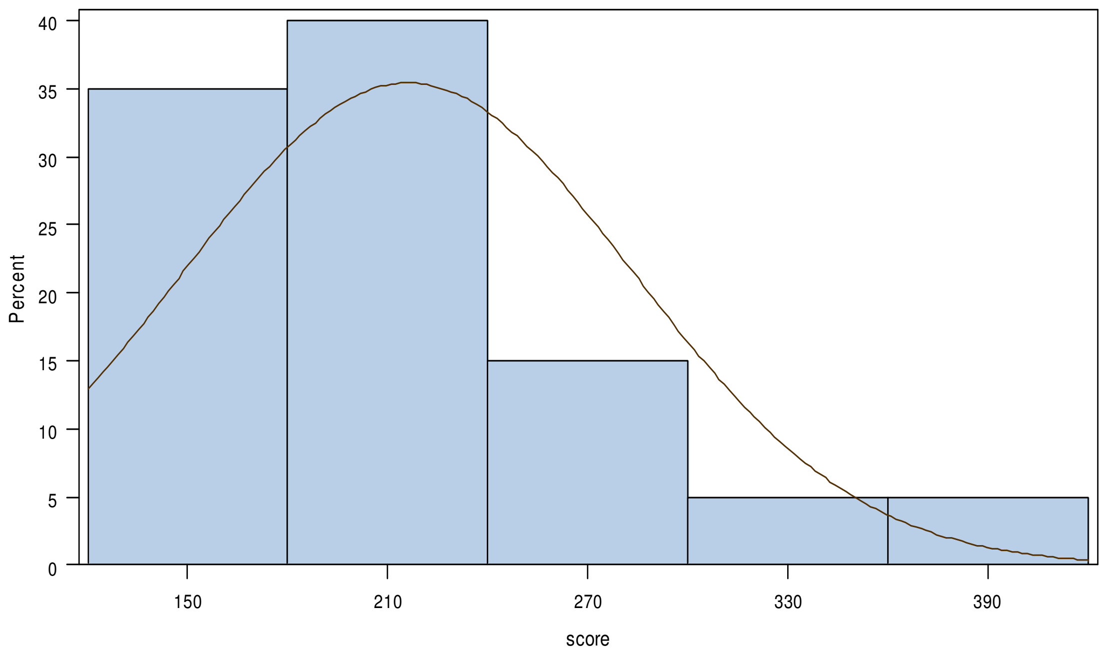

Figure A.7: Histogram of mean task completion time

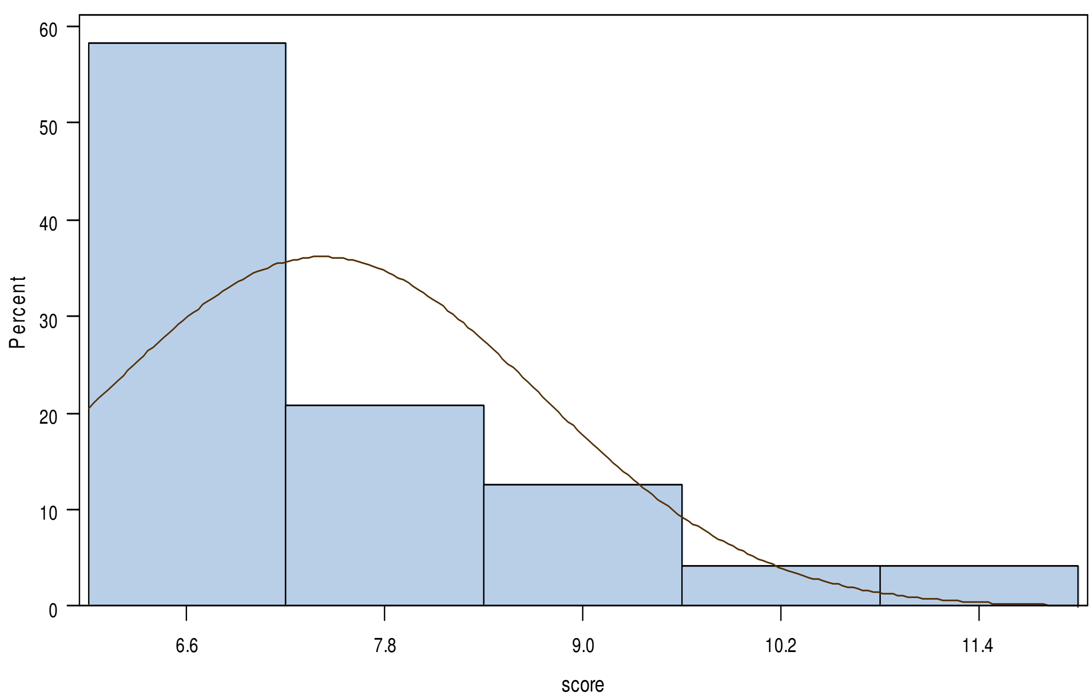

Figure A.8: Histogram of mean # of scoops

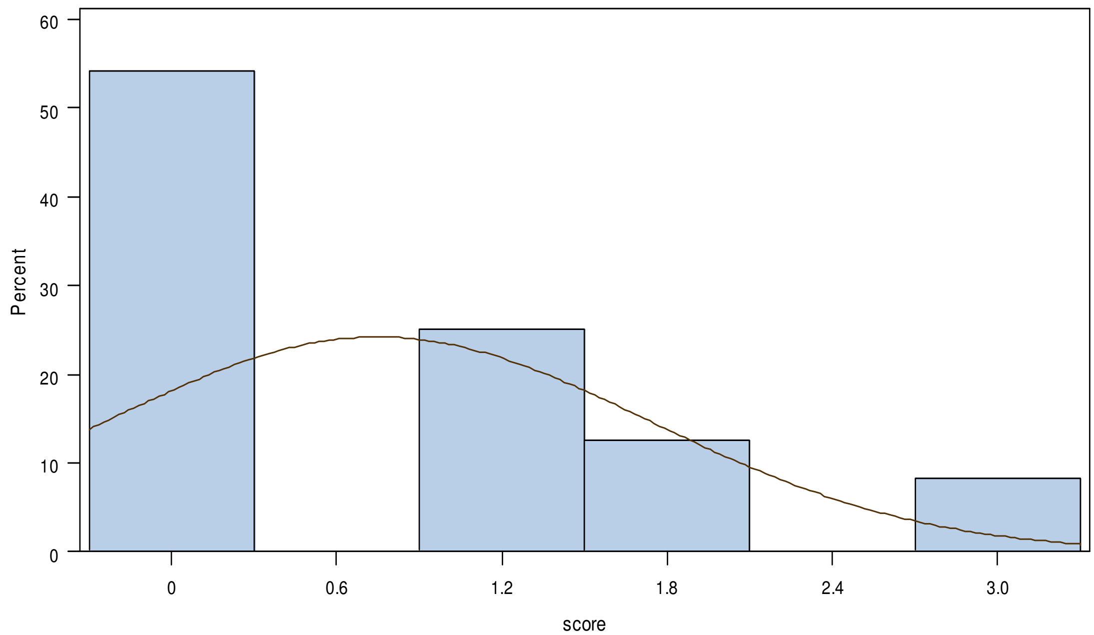

Figure A.9: Histogram of mean # of drops

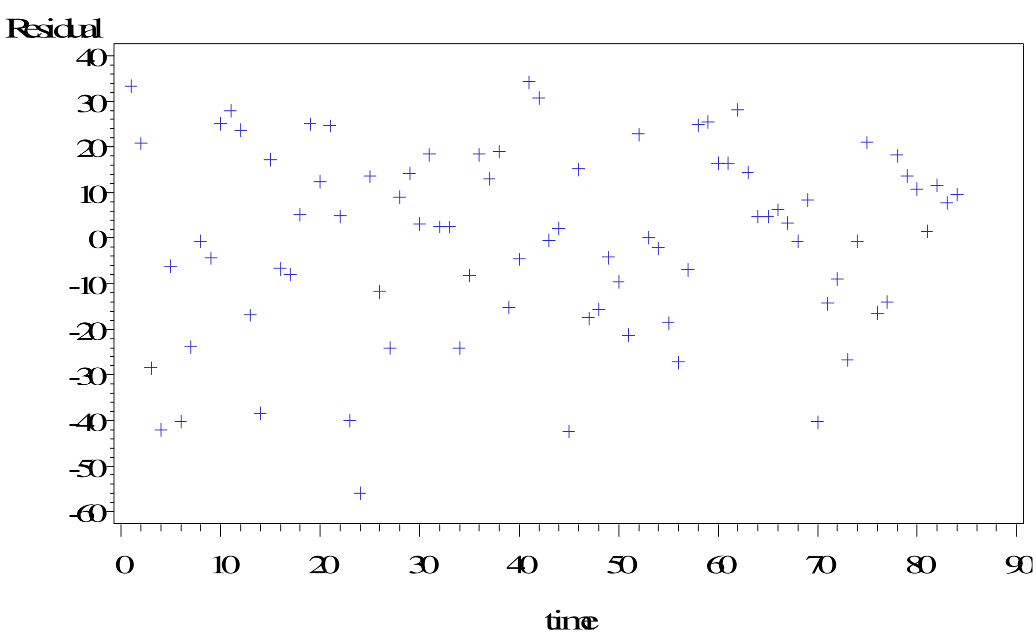

Figure a.10: Residual plot for No, Low, Medium and High Force Feedback

Appendix B

Table B.1: Task completion time for each participant

Subject	Task Completion Time (s)/Bin			
	No Feedback	**Low Feedback**	**Med. Feedback**	**High Feedback**
Subj1	132.0	138.0	100.5	108.8
Subj2	221.8	119.1	99.3	89.7
Subj3	466.4	271.7	179.6	249.4
Subj4	175.2	178.4	201.4	171.6
Subj5	217.8	133.2	175	222
Subj6	186.3	119.4	99.3	94.4
Subj7	125.5	128.4	137.9	104
Subj8	123.7	262.6	126.6	145.3
Subj9	161.1	217.0	196.2	184.5
Subj10	120.8	67.4	101.6	108
Subj11	197.8	91.2	123.3	113.6
Subj12	199.4	191.9	136.4	174.2
Subj13	253.2	119.2	140.2	112
Subj14	257.4	246.3	213.6	349.8
Subj15	320.8	268	258.9	172
Subj16	150.9	173.4	106.7	154.6
Subj17	292.4	263.6	252.8	211.8
Subj18	242.7	161.1	142.3	164.5
Subj19	202.0	100.5	66.9	98.9
Subj20	232.0	160.1	152.1	162.6

Table B.2: Number of scoops/bin for each participant

	Number of Scoops/Bin			
	No Feedback	**Low Feedback**	**Medium Feedback**	**High Feedback**
Subj1	6	6	5	6
Subj2	9	6	6	5
Subj3	10	8	6	7
Subj4	8	7	8	7

Table B.2: Number of scoops/bin for each participant *(cont.)*

	Number of Scoops/Bin			
	No Feedback	Low Feedback	Medium Feedback	High Feedback
Subj5	9	6	6	8
Subj6	7	5	5	5
Subj7	7	6	7	5
Subj8	9	9	6	6
Subj9	5	6	7	6
Subj10	8	8	10	11
Subj11	6	5	6	6
Subj12	10	7	5	7
Subj13	7	5	6	5
Subj14	6	6	7	7
Subj15	11	7	6	6
Subj16	6	5	6	6
Subj17	7	6	5	5
Subj18	6	6	6	6
Subj19	5	6	6	6
Subj20	7	6	5	6

Table B.3: Number of drops/bin for each participant

	Number of Drops/Bin			
	No Feedback	Low Feedback	Medium Feedback	High Feedback
Subj1	0	0	0	0
Subj2	0	0	0	0
Subj3	1	0	0	1
Subj4	1	2	1	0
Subj5	1	0	0	0
Subj6	0	0	1	0
Subj7	0	0	0	0
Subj8	0	0	0	0
Subj9	0	0	0	0

Table B.3: Number of drops/bin for each participant *(cont.)*

	Number of Drops/Bin			
	No Feedback	Low Feedback	Medium Feedback	High Feedback
Subj10	0	0	0	0
Subj11	0	0	0	0
Subj12	0	1	0	1
Subj13	1	0	0	0
Subj14	0	1	0	0
Subj15	1	0	0	0
Subj16	0	0	0	0
Subj17	0	0	0	0
Subj18	0	0	0	0
Subj19	0	0	0	0
Subj20	0	0	0	1

Appendix C

1. Task Analysis of excavator operation using GOMS (Goals, Operators, Methods and Selection) Approach.

Steps:

1. Data collection
   - Interviews were conducted and subject matter experts (SME's) were observed to gain in-depth understanding of excavation tasks
2. Definition of task under analysis
   - To conduct a task analysis of excavation task using the haptic controlled excavator
3. The overall goal of the task analysis
   - Use the haptic control excavator to dig a 6ft trench
4. Sub-tasks
   i. Start the excavator
   ii. Calibrate the haptic device (stylus)
   iii. Move the excavator to the desired work location
   iv. Position boom and bucket
   v. Scoop dirt/soil
   vi. Move dirt/soil to desired location and unload/release
   vii. Shut down excavator

5. Decomposing the sub-tasks into actions/operations

i. **Goal: Start Excavator**
Method for Goal: Start Excavator
Step 1: Stretch arm to reach starter/ignition key
Step 2: Turn starter ignition key
Step 3: Turn head
Step 4: Look at control panel (indicators/gauges)
Step 5: Recall that when excavator is turned on, *power on indicator* turns red
Step 6: Verify that power on indicator is lit/illuminated
Step 7: Return with goal accomplished

ii. **Goal: Calibrate the haptic device**
Method for Goal: Calibrate the haptic device
Step 1: Stretch arm to reach stylus
Step 2: Grab haptic device (Stylus)
Step 3: Recall that when stylus is moved up, the boom moves up
Step 4: Move stylus up
Step 5: Verify that the boom assembly moves up
Step 6: Recall that when stylus is moved down, the boom moves down
Step 7: Move stylus down
Step 8: Verify that the boom assembly moves down
Step 9: Recall that when stylus is moved left, the boom moves left
Step 10: Move stylus to the left
Step 11: Verify that the boom assembly moves left
Step 12: Recall that when stylus is moved to the right, the boom moves to the right
Step 13: Move stylus to the right
Step 14: Verify that the boom assembly moves to the right
Step 15: Recall that rotating the stylus clockwise and anticlockwise, opens and closes the bucket
Step 16: Rotate stylus clockwise the anticlockwise
Step 17: Verify that bucket opens and closes
Step 18: Return with goal accomplished

iii. **Goal: Move excavator to desired work location**
Method for Goal: Move the excavator to work location
Step 1: Stretch to reach lever
Step 2: Select forward/ backward lever
Step 3: Extend foot
Step 4: Press foot on travel/accelerator pedal
Step 5: Move excavator to desired work location
Step 6: Release foot from the travel/accelerator pedal

Step 7: Release forward/ backward lever
Step 8: Return with goal accomplished

***Selection rule* set for goal: Position boom and bucket**

If operator is inexperienced, then accomplish goal by using sequential positioning technique.

If operator is experienced, then use simultaneous positioning technique to accomplish goal.

iv. **Goal: Position boom and bucket using sequential positioning technique**

Method for Goal: Position boom and bucket using sequential positioning technique

Step 1: Stretch arm to reach arm controller/lever
Step 2: Grab swing arm controller/lever
Step 3: Extend foot
Step 4: Press foot on travel pedal
Step 5: Tilt/turn head
Step 6: Adjust swing arm controller/lever
Step 7: Position boom and bucket assembly
Step 8: Verify boom and bucket are at the desired position
Step 9: Release foot from pedal
Step 10: Release swing arm controller
Step 11: Return with goal accomplished

Goal: Position boom and bucket using simultaneous positioning technique

Method for Goal: Position boom and bucket using simultaneous positioning technique

Step 1: Reach arm
Step 2: Grab swing arm controller and simultaneously extend foot
Step 3: Press foot on travel pedal
Step 4: Tilt head
Step 5: Position excavator
Step 6: Verify boom and bucket are at the desired position
Step 7: Release foot and swing arm controller
Step 8: Return with goal accomplished

v. **Goal: Scoop dirt/soil**

Method for goal: Scoop dirt/soil

Step 1: Reach arm
Step 2: Grab stylus
Step 3: Recall that when stylus is pushed backwards (away from operator), the boom/arm assembly extends
Step 4: Push stylus backwards (away from operator) and up
Step 5: Verify that boom/arm assembly of excavator is extended
Step 6: Hold stylus steady and rotate in anticlockwise direction
Step 7: Recall that rotating the stylus anticlockwise opens the bucket

Step 8 Verify that the bucket is opened and the boom/arm assembly is extended
Step 9: Pull the stylus down and forward (towards operator)
Step 10: Verify that boom and bucket assembly is lowered
Step 11: Rotate stylus in clockwise direction and pull forward
Step 12: Verify that soil is scooped into bucket
Step 13: Move stylus up
Step 14: Verify that bucket and its content is above ground level
Step 15: Return with goal accomplished

vi. Goal: Move soil to desired location and unload/release

Method for Goal: Move dirt/soil to desired location and unload/release
Step 1: Hold stylus ready and move to either left or right
Step 2: Verify that boom assembly has moved to the left or right
Step 3: Stop when bucket reaches the desired location
Step 4: Rotate the stylus in anticlockwise direction
Step 5: Recall that anticlockwise rotation opens the bucket
Step 6: Verify that bucket opens and releases content
Step 7: Move stylus left or right
Step 8: Verify that the boom/bucket assembly moves back to the work area
Step 9: Return with goal accomplished

vii. Goal: Shut down excavator

Method for goal: Shut down excavator
Step 1: Reach with arm
Step 2: Select forward/backward lever
Step 3: Extend foot
Step 4: Press foot on travel/accelerator pedal
Step 5: Verify that excavator has moved to desired location
Step 6: Release hands and foot from pedal
Step 7: Reach hands
Step 8: Grab ignition key
Step 9: Turn ignition key
Step 10: Recall that when the ignition key is turned off, the ‘power on’ illumination goes off
Step 11: Verify that power has been turned off
Step 12: Return with goal accomplished