2013

Cooperative Interactive Distributed Guidance on Mobile Devices

Gregory Burnett
Wright State University

COOPERATIVE INTERACTIVE DISTRIBUTED GUIDANCE ON MOBILE DEVICES

A dissertation submitted in partial fulfillment of the
Requirement for the degree of
Doctor of Philosophy

By

GREGORY BURNETT
M.S., Wright State University, 2008

2013
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

April 12, 2013

I HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER MY SUPERVISION BY Gregory Burnett ENTITLED Cooperative Interactive Distributed Guidance on Mobile Devices BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF Doctor of Philosophy.

Thomas Wischgoll, Ph.D.
Dissertation Director

Arthur Goshtasby, Ph.D.
Director, Computer Engineering and Computer Science Ph.D. Program

R. William Ayres, Ph.D.
Interim Dean, Graduate School

Committee on
Final Examination

Thomas Wischgoll, Ph.D.

Arthur Goshtasby, Ph.D.

Yong Pei, Ph.D.

Gregory Funke, Ph.D.

ABSTRACT

Burnett, Gregory Michael Ph.D., Computer Science and Engineering Ph.D. program, Wright State University, 2013.Cooperative Interactive Distributed Guidance on Mobile Devices.

Mobiles device are quickly becoming an indispensable part of our society. Equipped with numerous communication capabilities, they are increasingly being examined as potential tools for civilian and military usage to aide in distributed remote collaboration for dynamic decision making and physical task completion. With an ever growing mobile workforce, the need for remote assistance in aiding field workers who are confronted with situations outside their expertise certainly increases. Enhanced capabilities in using mobile devices could significantly improve numerous components of a task's completion (i.e. accuracy, timing, etc.). This dissertation considers the design of mobile implementation of technology and communication capabilities to support interactive collaboration between distributed team members. Specifically, this body of research seeks to explore and understand how various multimodal remote assistances affect both the human user's performance and the mobile device's effectiveness when used during cooperative tasks. Additionally, power effects are additionally studied to assess the energy demands on a mobile device supporting multimodal communication. In a series of applied experiments and demonstrations, the effectiveness of a mobile device facilitating multimodal collaboration is analyzed through both empirical data collection and subjective exploration. The utility of the mobile interactive system and its configurations are examined to assess the impact on distributed task performance and collaborative dialogue between pairs. The dissertation formulates and defends an argument that multimodal communication capabilities should be incorporated into mobile communication channels to provide collaborating partners salient perspectives with a goal of reaching a mutual understanding of task procedures. The body of research discusses the findings of this investigation and highlight these findings they may influence future mobile research seeking to enhance interactive distributed guidance.

TABLE OF CONTENT

LIST OF FIGURES

1. INTRODUCTION

Mobiles device are quickly becoming an indispensable part of our society. Equipped with numerous communication capabilities, they are increasingly being examined as potential tools for civilian and military usage to aide in distributed remote collaboration for dynamic decision making and physical task completion. Remote collaboration on physical tasks is defined by Kraut et al. (2003) to be: “A general class of ‘mentoring’ collaborative physical task, in which one person directly manipulates objects with the guidance of one or more other people, who frequently have greater expertise about the task.” (p.16) The ability for mobile device users, hereafter referred to as Workers, to request non-collocated assistance from experts, hereafter referred to as Helpers, when confronted with situations that are outside their expertise could provide significantly improved results in terms of task completion, performance times, and accuracy, among other measures. Consider the following examples and effects of remote collaboration through mobile devices: survivability of time-critical casualties being attended to by in-field medics or first responders through the guidance of a remote surgeon; repair of machinery by the end user advised by a non-collocated mechanic; troubleshooting complex electronic systems by an untrained electrician under the expert guidance of highly trained electronics personnel.

Effectively relayed task knowledge is paramount to promoting efficient remote collaboration; however, to effectively collaborate, Clark and Brennan (1991) report there needs to be a mutual understanding between Helper and Worker to ensure common ground. This concept of common ground, or clarity of instructional directives, can be achieved through various modalities. Visual, auditory, and haptic modality information can be leveraged to provide easyto-process and environment appropriate interactive communication sessions exchanging perspectives and task information between Helper and Worker pairs. The very nature of mobile devices and user mobility further challenges effective remote collaboration. Workers' environmental contexts surrounding a cooperative task may change unexpectedly rendering certain modalities ineffective for receiving collaborative guidance. For example, a Worker and Helper's interactive communication may be disrupted if it occurs through the auditory modality and ambient noise levels are elevated to a point where it begins masking information exchanged between the Helper-Worker team. Moreover, consider a Helper and Worker remote collaboration leveraging visual information to provide directives for a cooperative task and the ambient lighting or visual demands of the environment changes, requiring reallocation of visual focus. Therefore special attention should be given to support multiple modalities in facilitating remote collaboration using mobile devices. Wickens and McCarley (2008) suggest that systems and interfaces that utilize multiple modalities are more advantageous to users than those that do not. As highlighted in the examples above, multimodal interfaces allow users to process different modality information concurrently for better cognitive understanding of the task. The ability to process multiple modalities concurrently fits well into the mobile domain and responsiveness required for mobile devices functioning in dynamic environments.

With the unprecedented growth of mobile devices and established mobile networks, the ability for on-the-move individuals seeking knowledge from a distributed source is becoming a reality. This collaborative communication between mobile users executing in unpredictable environments as well as working on diverse tasks warrants an assessment of both the mobile users' and mobile devices' performance. Mobile devices are equipped with numerous embedded communication capabilities that can support real-time remote collaboration, but at what cost? Brehmer et al. (1992) have investigated the effects of data delivery timeliness on the outcome of dynamic decision making situations. However, little is known about how mobile device characteristics affect a user's ability to complete a cooperative task under the remote guidance of a subject matter expert. Moreover, it is unknown how the user's performance is affected by communication through the multiple modalities (e.g. visual, auditory, and haptic) that mobile devices can leverage to convey data. Furthermore, whether the devices themselves can support the processing/power demands required to adequately enable these communication capabilities to function for the duration required to complete a cooperative task.

Cooperative tasks may require the mobile device to be interoperable with heterogeneous distributed systems through a variety of communication channels, allowing distributed individuals the ability to share real-time data and individuals' perspectives. The proliferation of mobile devices in today's society fosters countless interoperable collectives consisting of one-to-one, one-to-many, and many-to-one integrated dissemination and data processing. This interaction between mobile devices and users, while greatly beneficial, needs to be enabled efficiently to prevent over-stressing the platform's processing resources and drain the device's battery. Effectively capitalizing on mobile devices' characteristics is an attractive arena for distributed collaboration and peer-to-peer guidance scenarios. Sharing of real time mobile device data (e.g. audio, video, etc.) with physically separated computing platforms and individuals is of interest to those seeking instruction on procedures to fulfill a given objective in an effective and efficient manner.

This dissertation seeks to demonstrate the effects that modalities have on remote collaboration between distributed entities utilizing mobile devices with respect to human performance and power consumption. The research documents the design and implementation of an Android interactive application that leverages multimodal communication capabilities facilitating remote collaboration. Studies using this system are executed to highlight the human performance effects and mobile device utilization during performance of cooperative, distributed objectives. Additionally, an investigation is conducted focusing on power conservation as

Helpers and Workers negotiate and simultaneously monitor power consumption effects of active multimodal communication capabilities for the purpose of ensuring power consumption and battery life does not limit cooperative task performance.

Specific experiments demonstrate theoretical ideas that multiple modalities are more advantageous than unimodal interfaces when attempting to collaborate on complex distributed tasks (Wickens, 2008). Additionally, the amount of time required for the convergence on a mutual understanding between the cooperative pairs varies depending on the combinations of modalities leveraged in the communication exchange. Practical examples, especially important to the USAF, highlight the effects on user performance, confidence, and trust when collaborating with a non-collocated subject matter expert to execute cooperative tasks such as medical treatment, improvised explosive device disposal, and “find-fix-tag target of interest” scenarios.

2. BACKGROUND

The intent of this literature summary is to provide supporting background on the development of an effective multimodal mobile remote collaboration capability. The chapter begins by highlighting the need for effective communication between distributed collaborating pairs. This is followed by a review of research describing the effects of modality on communication in teams. As remote guidance techniques are still maturing for mobile devices, the background section will focus on a large contingent of computer supported cooperative work (CSCW) that has been executed on static workstations. Although certain communication capabilities and features described within can be leveraged, others are not suitable for the mobile domain. Additionally, a review of mobile power management and consumption techniques is described.

Throughout this review section, discussed distributed systems and research investigations strive to improve the interactions between cooperative pairs executing a collaborative physical task. A collaborative physical task is defined by Kraut, Fussell, and Siegel (2003) as: *"A general class of 'mentoring' collaborative physical tasks, in which one person directly manipulates objects with the guidance of one or more other people, who frequently have greater expertise about the task" (p.16).* Common classification descriptors for the individuals involved in distributed task are Helper and Worker. Helper refers to the individual with expert knowledge about a given task, and who provides instructional or directive information to a Worker. Worker refers to the individual applying the instructional or directive information to the local scene or workspace.

Figure 2.1: Helper/Worker collaboratively completing a physical task

This background review section will specifically look at communication, auditory sharing, video sharing, markup and annotation, and power management techniques.

Existing collaborative systems do not have the flexibility built into their communication link between Helpers/Workers to dynamically change resolution and fidelity of the information that is shared. Additionally, with existing CSCW systems, power consumption is not considered in the design of the systems as they rely on “unlimited” power (i.e. wall outlet plugs) supplying ample energy for the duration of their collaborative session. However, to effectively implement a collaborative interactive communication session onto a mobile platform, power and processing considerations are essential.

This dissertation’s research implementation allows for dynamic resolution and fidelity changes of disseminated multimodal information in a power conscience manner. An additional feature the newly designed system supports, and the current CSCW systems lack, is the ability for both Helper and Worker parties to monitor in real-time power levels and effects. This feature is a significant advantage incorporated in the new system over traditional static and emerging mobile collaborating interfaces. The implemented collaborative system is designed to collect and report to all parties the current state of charge for the mobile device, which is generally the weakest link in a distributed system due to its finite battery source. With knowledge of the current state and of how the various communication capabilities leveraged by the cooperative system effects power consumption, the interactive interface promotes real-time negotiation of data transmission that seeks to prolong the mobile devices' battery for the duration of the cooperative tasks. Additionally, a power estimator enables Helpers/Workers quick feedback on their changes in respect to power consumption and time extension. Another feature that distinguishes the developmental system from other collaborating systems is the ability for distributed entities to modify local settings of a mobile device remotely. This feature adds greater versatility for remote assistance, while mobile device users are engaged with the task's workspace. A non-exhaustive local list of settings that can be dynamically changed includes: display brightness, audio levels, interface control activation, disabling communication capabilities, and changes to amount and type of information disseminated.

The system demonstrated in this research also enables Workers greater control of received remote visual information over existing CSCW systems. Remote annotated image's transparency and orientation can be modified prior to fusing them with the current perspective. This capability has been demonstrated with limited functionality with the use of external projectors that render a scene on-top of the active workspace. The new collaborative mobile system allows remote collaboration anywhere at any time, supporting features and techniques that are self-contained internally to the mobile device without the use of external hardware. The system additionally equips the mobile user with the reactive ability to choose a video capture source. This is another feature that the system in this research supports that other CSCW systems generally exclude.

With the flexibility to capture the local workspace from different perspectives using available internal or external video capture devices, this system can scale to meet the mobile demands of countless users. Finally, the ability to hand hold the developmental system further distinguishes this remote guidance system from CSCW systems that have large footprints and are desktop bound.

The following sections will be a good reference when reading the experimentation and design sections.

2.1 Communications

Communication between distributed collaborating pairs working together on a remote objective is paramount to the successful completion of a physical task. Communication can be achieved through various modalities: auditory, visual, and/or haptic means may be employed to relay information between Helper and Worker, each of which may play a useful role in successful communications (e.g., Gergle et al., 2004). Regardless of modality, effective cooperative communication leads to a common understanding of the procedures and components involved in a task at hand. This common understanding is often referred to as communication grounding. Cognitive theorists Clark and Brennan (1991) explain communication grounding as a collective process by which Helper/Worker pairs try to reach a mutual belief or knowledge through progressive conversation consisting of *Presentation* and *Acceptance* phases. In a physical task, the Helper typically conveys the *presentation* phrase and the Worker responds with an *acceptance* phrase as they work towards the completion of a task. For example:

Helper: The next structure piece is a blue four-by-four block. (*Presentation Phrase)*

Worker: Ok, got it (*Acceptance Phrase)*

Although in the cases where the Worker needs additional information or needs further clarification, they can introduce a *presentation* phrase to the Helper and a corresponding *acceptance* phrase is acknowledged with a new *presentation* phrase. For example:

Helper:	Place the blue block on top of the red block.	(*Presentation Phrase)*
Worker:	Um, which red block?	(*Presentation Phrase)*
Helper:	The red block on the far right side of the structure.	(Acc/*Pres Phrase)*
Worker:	Roger that.	(*Acceptance Phrase)*

This dynamic dialogue exchange between Helper/Worker tends to conform to the principal of *least collaborative effort*. According to Clark and Wilkes-Gibbs (1986), the principal suggests that communicating pairs selectively utilize the minimum amount of information/effort to relate their directives or understanding during remote collaboration tasks. The use of multiple modalities may facilitate team performance (and reduce the influence of the least collaborative effort principal) by affording communicating pairs with additional channels with which to communicate, resulting in less time and resources expended and fewer errors (Clark & Krych, 2004).

2.2 Sharing of Auditory Information

In auditory sharing, the "sender" relays information through acoustic signals, which the "receiver" must then decode and interpret (e.g., Buck & VanLear, 2002). A traditional example is the telephone, where verbal dialogue exchange is conducted in the absence of visual or haptic data to express the points or perspectives of the speaker. Auditory information when dealing with remote collaboration between a Helper and a Worker can be categorized into two overarching modes: input and output. These modes can be introduced and processed *cotemporality*, *simultaneity*, and/or *sequentiality* by collaborating pairs during a communication exchange leveraging auditory information (Clark & Brennan, 1991).

As an input mechanism, auditory interaction enables hands-free control of applications/features without necessarily drawing focus away from the task at hand (Smailagic, 1998). This is critical to remote collaboration on physical task as situation awareness (i.e., the current state of the task) and the ability to use one's hands, generally the primary tool used in a task, is important to task completion. Additionally, auditory input does not require visual or physical contact with a device, enhancing communication convenience for remote collaboration (Zaykovskiy, 2006). Auditory signals can additionally serve as input trigger mechanisms to adjust information portrayal of data to a mobile device user in a mobile context (Haggon, 2009). For example, if ambient auditory levels surrounding a mobile device reach a threshold, where audio data may be masked, then auditory information may be better represented in textual form. Moreover, auditory signals can be used as an input source in Speech-to-Text and Speech-to-Control capabilities. For example, Schuster's (2010) Voice Search application enables mobile users to use speech input to conduct search queries instead of having to physically interface with a mobile device's keyboard. This feature could prove useful when Workers need to search a document provided by a Helper for key procedural steps on a task. Ballinger et al. (2010) focused on the processing aspects of speech-to-text with an on-demand speech interpolation finite state transducer (FST) for improved mobile speech recognition performance and control. Their on-demand FST calculates interpolation weights for input utterances from several *n*-gram language models resulting in an 11.2% reduction in word error rate. This would lessen the processing and power consumption associated with speech-to-text and speech-to-control as user would not be have to reiterate verbal inputs.

As an Output mechanism, the most common auditory sharing communication technique is speech-to-speech. Auditory collaboration between a Helper and a Worker is performed through the exchange of informative verbal phrases (i.e., descriptions, directives, acknowledgments, request for clarification, etc.) to arrive at a mutual understanding. For example:

Helper: "Place the blue block on top of the red block"

Worker: "Got it"

Gale (1990) found in specific applications that access to a good-quality, full duplex auditory communication channel resulted in faster team task completion times than using audio and video. Auditory signals can additionally serve as an output means through devices that convert textual information into speech, or by translating and broadcasting a spoken message in a user-selected language. Furthermore, auditory signals can serve as an output source when used as an alert or notification. As an alert/notification mechanism, auditory sharing can enable users to retain visual focus on their task while processing the auditory information (Pirhonen, 2002). This is key for mobile remote collaboration as Workers may have to divide their cognitive attention across several events simultaneous during task execution (i.e. the task and the environment surrounding the task). Moreover, the use of auditory signals is an effective way to trigger or focus one's attention to a particular event or a status change (Gaver, 1997).

A technical feature implemented in this research, but often overlooked in the implementation of auditory information sharing in other collaborative systems, is the ability to adjust audio quality and output levels dynamically. The captured auditory sharing research lacks the adjustability of real-time audio properties in bandwidth limited and power constraints scenarios. These features when operated on mobile devices in mobile use cases could prove beneficial for collaborating teams. The implementation needs to scale to the demands and limitations of the mobile user and surrounding environments of the mobile device and its current operations. Therefore, auditory information sharing implementation for this dissertation research will be done in a power conscience way that permits either cooperative team member to adjust the audio quality in real-time. Additionally, the presentation of auditory information (i.e., playing audio signals on speakers) could consume precious power on mobile devices. The ability to dynamically adjust the loudness of auditory information is implemented in a way that allows remote parties to adjust output levels.

2.3 Sharing of Visual Information

During remote collaboration, the ability to exchange information using a visual medium has been shown to decrease task completion time, improve task accuracy, reduce the amount of verbal information exchanged, and increase confidence and trust in Helper/Worker teams (Fussell et al., 2000; Gergle et al., 2004; Kirk et al., 2005; Kraut et al., 2002; Kraut et al., 2003). Researchers have explored several methods for sharing visual information between Helpers and Workers. Examples include utilization of static cameras to monitor a Helper/Worker's immediate workspaces, using a head/helmet mounted camera that captures the immediate field of view of the Worker/Helper, and a stationary, mounted camera with motorized range of motion controlled by the Worker/Helper.

Shared visual information can serve four supporting roles in remote collaboration: awareness, detection, confirmation, and adaptation. Awareness of the cooperative task's current state can be enhanced by sharing visual information between distributed cooperative partners. For effective remote assistance, Orr et al. (1996) argue that Helpers must maintain consistent awareness of Workers' actions, the current state of the task, and the active workspace. Kraut et al. (2003) further articulate that collaborative awareness enables the Helper to assess ongoing task progress/success, and determine what information is required to be presented next to the Worker.

Kuzuoka (1992) additionally suggests that congruence between the focus of a shared visual display and Worker activity improves Helper situation awareness, resulting in better guidance from the Helper to the Worker. “Visual information can give collaborators an up-to-date view of the state of the task. Additionally, it provides evidence about a partner’s level of understanding of the language that is being used for coordination.” (Kraut et al., 2002, p. 31).

Adaptation of communication between collaborating pairs is positively affected by sharing visual information, especially when cooperative tasks are “visually complex, dynamically changing or when the objects in the display are difficult to describe linguistically” (Gergle, 2005, p.1117). A shared visual display may be more efficient than an auditory-only communication channel as Helpers can leverage the visual information rather than explicitly questioning Worker understanding. For example, Kraut et al. (2003) demonstrated that Helpers elaborated more and provided more detailed instructions when they were able to monitor a Worker’s comprehension with a shared visual display. Isaacs et al. (1993, p.199) “found that, compared with auditory-only, a video channel adds or improves to show understanding, forecast responses, give non-verbal information, enhance verbal descriptors, [and] manage pauses.” Kraut et al. (1996) report that the manner in which collaborating pairs coordinated guidance varied when they leveraged a shared visual space. The authors highlight that assistance was “more proactive and coordination was less explicit when the pairs had video connections” (p.57).

Detection of errors and prevention of compounding or nested errors are lessened when shared visual information is disseminated between Helper and Worker. Gergle (2005) found that “pairs are able to detect errors earlier on in the course of their work and remedy the situation in a timely fashion before their actions become nested and they need to revert through several previous task states in order to fix any problems” (p. 1117). Kraut et al. (2003) also noted that with the advent of visual information sharing between Helper and Worker, Helpers can “determine if clarification or expansion of the instruction are required.” (p. 18). For example, if a Worker makes a mistake, the Helper can “interject a comment to correct” (p. 18) the action. In remote collaboration, the prevention or mitigation of errors could significantly affect task performance and overall outcome.

Confirmation of a remote directive can be witnessed through shared visual information. By observing the actions of the Worker, the Helper can recognize when the Worker is confused and does not comprehend the instructional guidance, or when the worker does not understand the general task (Brennan, 2004). Gergle et al. (2004) further support that visual information serves as an important “feedback loop to get verification both that an instruction had been heard and that it had the intended effects” (p. 489). Kraut et al. (1996) argued that “when the worker and expert share a visual workspace, the expert can receive feedback from the task itself to precisely time when he gives instructions and which instructions to give” (p. 58). For time sensitive cooperative tasks, confirmation that a procedural step is correctly accomplished can efficiently progress the task toward completion. This was illustrated by Kraut et al. (2002), who stated that “when the Director [Helper] could see what the Matcher [Worker] was doing, the pair was substantially faster, in part because the pair could precisely time their words to the actions they were performing” (p. 32).

Existing visual sharing collaborative systems utilize a combination of various video capture devices (VCD) and display components. Workers wear a VCD tied to their head and a head mounted display (HMD), which renders images or video from a remote Helper. Fussell et al. (2000), Kraut et al (1996), Kuzuoka et al. (1994) used this configuration, where Workers wore a small CCD camera mounted to their head and a low resolution (480x600 pixel) HMD. Both devices were tethered to a nearby computer that supplied power and data transmission of the visual information through a hard line network. Technical shortcomings of these systems that limit their mobility and deployment is that the VCD and display are not integrated, visual data dissemination is not wireless, VCD and display worn by the Worker does not support remote changes, and power consumption is not considered.

In contrast to the above systems, the visual information sharing implementation for this dissertation research is done in a power conscience manner that permits either cooperative team member to adjust the video quality in real-time. Additionally, the acquisition and presentation of visual information (i.e., capturing frequency of video and display's brightness) could consume precious power on mobile devices. Therefore, the ability to dynamically adjust the brightness of the Worker's display and the camera's captured frames per second was implemented in a way that allows either remote or local parties to modify visual dissemination settings. Also, in this implementation, visual information is transmitted wirelessly.

2.4 Sharing of Markup Annotations and Fusion with Active Workspaces

A compliment to visual information sharing is the ability to add graphical information in the form of markup annotation or gestures to the shared visual information or active workspace between Helper and Worker. Communication between cooperative pairs is often facilitated with gestures that highlight an object of interest, drawing attention to a particular region, or illustrating the use of an item. Moreover, markups and gestures can simplify the spoken dialogue between Helper and Worker through the use of pronouns such as "this one" and "over here" while highlighting items or regions within the active workspace. The ability to add information to an active workspace has been shown to enhance understanding and task performance (Ou et al., 2003; Kirk et al., 2007; Stevenson et al., 2008).

Ou and colleagues' (2003) *Drawing Over Video Environment* (DOVE) remote collaborative system allowed a "remote helper to draw on a video feed of a workspace as he/she provides task instructions." (p. 100). The DOVE system supported both free-form annotation as well as gesture fitting recognition to generate a markup perspective shared with the Worker. Results from their research suggest that markup capability "significantly reduces performance time compared to camera alone." (p. 248).

Kirk et al. (2007) designed a collaborative video/audio environment that sought to address mixed reality ecology by conjoining two separate but similar workspaces into one hybrid workspace. The interactive system overlaid video-captured gestures and workspace elements of the Helper onto the active workspace of the Worker through the use of projectors. Creating a linked collaborative workspace, the Helper could direct the Worker's actions through the use of simple hand gestures, illustrated marks using a pen, and auditory commands. Their results showed that task completion time was shorter and error rates were reduced when the Helper used the combination of auditory commands and gestures.

Stevenson and colleagues' (2008) research utilized a combination of "on-video" and "in-workspace" annotation capability, where a remote Helper could use illustrated guidance to direct the action of the Worker. The use of annotation techniques reduced the spoken instructions into "spoken fragments like 'in', 'out', 'around', and 'here' as they drew" (p. 38) their remote directives. Their results showed that the utility of annotation affected verbal communication allowing collaborating members to be more efficient and able to using verbal shortcuts in distributed communication.

Technical shortcomings of existing remote annotation sharing systems are the limited- or non-existence of bi-directional annotation between Workers and Helpers. Because of this, the Workers have no control over how the annotation is projected onto their display or workspace (for example: placements, orientation, or intensity). Additionally, existing systems do not permit Workers to easily “look through” remote annotations to apply markup information to the local workspace. Therefore, markup annotation sharing implementation for this dissertation research will address these shortcomings. The capability for bi-directional markup generation is implemented in this research. Here, controls will be given to the Worker enabling them to adjust the presentation of remote markups that best accommodates their current activities, as well as the ability to adjust the remote annotation’s transparency.

2. 5 Haptic Sharing

The use of haptic information is an evolving modality in mobile devices where vibrating tactile sensors are used to relate instructions, notification, and other relevant information (Luk et al., 2006). When used in visually and auditorily distracting environments (e.g. urban cities and subways), haptic sensors have been shown to reduce the cognitive workload of interacting with mobile devices in the retrieval of data through the sense of touch (Oulasvirta et al., 2005). Tactile icons, or “tactons” (e.g., Brewster & Brown, 2004), can provide information through the sense of touch, and can be represented to the user by manipulating several parameters, including amplitude, frequency, duration, and waveform. It’s been suggested that tactons can improve interaction in various mobile contexts and usages (e.g., Brewster & Brown, 2004). Additionally, haptic modality use can potentially offload display communications, and increase perceptual bandwidth available for mobile information interactions (Chang & O’Sullivan, 2005). Haptic information sharing can be grouped into two main categories: feedback/notification and information portrayal.

Feedback/notifications can be enhanced by tactile presentation through the use of vibrations to indicate various conditions such as alarms, alerts, or incoming calls. Tactile vibrations have been used to provide status information on mobile processes status such as when messages have been sent or arrived from/to a mobile device. Haptic feedback was investigated by Hoggan et al. (2009) as a means for mobile users to perform messaging tasks while riding on a subway. The authors determined that haptic feedback was effective at vibration levels below 9.18 g/s. Additionally, research efforts are leveraging haptic sensors in mobile applications to present and capture input data for interactions between users (Chang & O'Sullivan, 2005; Heo & Lee, 2011; Linjama & Kaaresoja, 2004) A limitation with haptic information sharing is the increase in power consumption associated with its use and potential reduced usefulness when the mobile context features vibration.

Information can be relayed to mobile users in the form of haptic pulses. Similar to brail for the blind, haptic pulses can be presented as unique tactile stimuli associated with functional meaning. Luk et al. (2006) describe a hardware concept that can be added to mobile devices that can produce a wide range of tactile output as tactons. MacLean and Enriquez (2003) used haptic icons to represent abstract messages to mobile users to describe an object's and event's current state, context, or function. Their research suggests that users were able to learn and interpret a small set of tactile stimuli; however their recognition performance decreased as users divided their cognitive resources to interact with their surroundings.

2. 6 Power

Mobile devices have always been limited by their batteries, i.e., by their finite storage capacities and the power consumption rates of the devices. Satyanarayanan (1996) highlights battery consumption among the major challenges in mobile computing, along with processing and connectivity. Advancements in mobile processors (e.g. dual cores) and network infrastructures (e.g. 4G) have greatly improved the processing and connectivity aspects of mobile computing across manufacturers, however power consumption and management techniques are still not standardized. Mobile devices will always have a finite energy source, as a battery's size and weight are constrained by the device it powers (Carroll, 2010). Therefore an understanding of desired features and their associated power consumption is paramount to fostering smarter power management that ultimately prolongs the battery run-time (Kjaergaard & Blunck, 2012). In regards to remote collaboration between Helper and Worker, power consumption awareness and sustainability of the communication link enabled by a mobile device until the completion of a cooperative task is critical.

In the pursuit of power conservation, power measurements of mobile devices have been researched from numerous perspectives, identifying power consumption models, studying empirical findings of results, and explaining emerging measurement techniques. Kravets and Krishnan (1998) investigated the technique of managing the cycles of the transport layer (e.g. suspending/resuming) of the mobile host's communication to reduce power consumption. Their results showed energy consumption savings of approximately 6-9%; however, this introduced latency in the bidirectional incoming and outgoing data. When dealing with remote collaboration and guidance, latency of information in one direction or both may prove to be acceptable in some tasks, but may be detrimental in others. Kremer et al. (2001) evaluated the energy consumption savings that offloading complex calculations to another system connected via a wireless network afforded. Their initial finding showed “in some cases up to one order of magnitude [savings], depending on the selected characteristics of the mobile device, remote host, and wireless network” (p. 1).

The dynamic time requirement of distributed collaborating tasks between Helper and Worker demands power consumption of both hardware and software to be optimized. Kjaergaard and Blunck (2012) suggest that, to efficiently minimize the cumulative power consumption and promote improved power conservation, knowledge of all specific communication features power affects is needed. To obtain power consumption information, power profiling can be used to measure the total power consumption of an operation or process and can be done either through hardware or software means. Dong and Zhong (2011) performed a comparison of power models constructed through internal battery profiling (software) versus external equipment (hardware) and found the resulting profiles only differed marginally. Flinn and Satyanarayanan (1999) sought to measure the hardware and software power consumption contributions that individual applications consumed while attempting to meet a user-specified battery duration. Their approach used an in-line hardware multimeter to determine the power usage of isolated hardware and software components as they operated concurrently. With accurate profiles they were able to show energy reduction greater than 7% as they used the profiles to adjust fidelity and resolutions of mobile capabilities. More recently, as mobile devices are experiencing frequent software upgrades to applications, drivers, and operating systems, software power profiling affords better scalability than hardware profiling. Software profiling can be executed on-demand or following an upgrade more easily than hardware profiling, as no external equipment is necessary to construct new power models (Kjaergaard & Blunck, 2012).

Knowing the power consumption profiles associated with mobile device communication capabilities allow researchers and software developers alike to support application-awareness for improve power efficiency and duration. Rao et al. (2003) reports that by equipping mobile users with knowledge of the tradeoffs in performance and battery life, users can actively participate in power consumption management to meet their needs. In regards to remote collaboration, if team members were armed with information about mobile communication capabilities and their respective power consumption effects, then teams could adjust their communication strategies to ensure that battery life survives for the duration of required interaction. This dissertation implements a power monitoring capability that shares a Worker's mobile device power state and current power consumption rates to remote Helpers.

3. METHODOLOGY AND DESIGN

To research user performance, cooperative guidance, and mobile device capabilities in processing modalities in diverse scenarios, a theory-supported research project was conducted leveraging newly developed Android software, rigorous performance metrics, and relevant use-case experiments. To gain awareness of how multi-modal communication affects remote collaboration and physical task completion, an experimental mobile application was implemented. This mobile application was developed to run on the Android operating system. Android was selected because of the military's interest in using the Android OS to host various on-the-move capabilities due to its open source nature and flexibility in running third party software. The mobile device chosen to evaluate the effects of the modalities was the Samsung Galaxy Tablet, however the mobile application can be run on any Android supported device. The following chapter highlights the implementation of the collaborative communication system that facilitates the connection of a mobile device user termed the Worker, to a remote expert termed the Helper. The connection functions similarly to a client/server distributed system, although each side can independently initiate the various communication techniques as they see fit. Moreover, the mobile application permits several simultaneous connections with remote entities as it supports multicast communications in receiving and disseminating cooperative data between Helper and Worker roles. The mobile application was tested in interactive trials where the Worker was in communication with a Helper via WiFi connectivity during completion of specific task objectives.

The prototype development described herein is part of an on-going research program that focuses on the design and development of advanced wearable interface technology for Battlefield Airmen. A user-centered design approach was employed with the explicit goal of designing multimodal, context-rich functionality into the mobile application to improve interactive collaboration of non-collocated parties. Drawing from documented related research approaches and other mobile application resources, a unique combination of the following capabilities were implemented in the hereafter described mobile application distinguishing it from tradition CSCW systems.

- Sharing live video of active Worker's workspace.
- Sharing full duplex audio between linked users.
- Supporting free form and predefined markup annotation.
- The ability to adjust transparency of overlaid markup images superimposed over live workspace view.
- Dynamic user configurable display modes representing adjacent and merged preview perspectives.

All the highlighted capabilities are describe in detail below.

3.1 Status Message

Across communication capabilities that are active during the collaborative session between a Worker and Helper, there is a constant status message transmitted between the pairs. This status message permits real-time negotiation as well as the ability for either the Worker or Helper to adjust the settings of communication and information rendering for the mobile device. This status message is a feature that allows the mobile application to provide improved communication and duration of interactive sessions compared to currently existing CSCW systems. The status message usage is highlighted in each of the following sections and its respective controls are discussed. Listed below are the contents of the status message:

X-Value: integer value of the starting position of the still image ($0 \leq X \geq 800$)
Y-Value: integer value of the starting position of the still image ($0 \leq Y \geq 600$)
Transparency Value: current transparency value of the still image ($0 \leq T \geq 255$)
Acknowledge Flag: Boolean flag indicating messages have been processed by Worker (T or F)
Full screen Flag: Boolean flag indicating the current preview mode used by Worker (T or F)
Shared Frames: integer value of the shared video frames per second (30, 16, or 6)
Battery Charge: float value indicating the current charge of the mobile device (100.0 – 0.0)

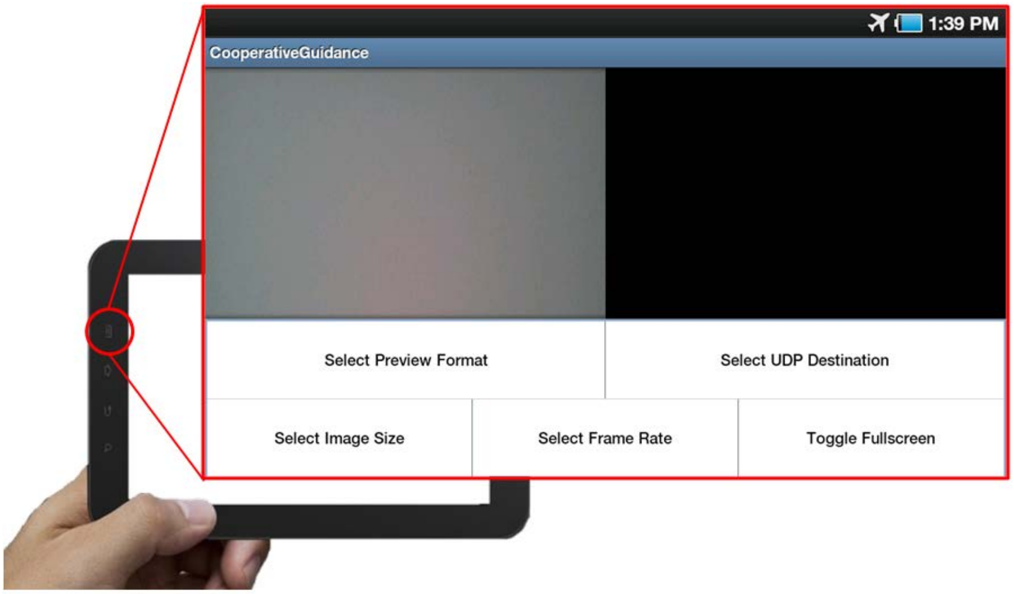


Figure 3.1: Mobile Application Control Menu

3.2Control Menu

Complimentary to the Status Message, a control menu was designed as part of the mobile application to adjust various settings of the communication link during run-time locally by the Worker. The on-demand options foster improved power conservation for the capturing, packaging and transmission of the multimodal information shared between Worker and Helper. Figure 3.1 depicts the menu and its corresponding options that are presented to the Worker upon selecting the menu button on the mobile device. The menu enables the modification of: image format, image size, frame rate of image transmission, destination of TCP/IP remote collaborators, and mobile display mode. The resulting control and implementation of the menu options are described in the following sections.

3.3 Multimodal Communication Capabilities and Overarching Design

Drawing from related work, the use of multiple communication modalities has been shown to efficiently and effectively support remote collaboration on traditional CSCW systems. This dissertation explores and leverages a variety of multimodal communication capabilities shown to improve human performance in workstations and adapts them to work on a mobile device. The following sections discuss the design and implementation of a new Android multimodal mobile collaborative communication capability.

Figure 3.2: Mobile Collaborative Capability Overview

3.3.1 View Sharing

The “View Sharing” feature of this system provides the capability to capture and relay the perspectives of the Worker’s task environment to the Helper. Additionally it facilitates the receipt and display of remote visualization data from the Helper. The dissemination of the local

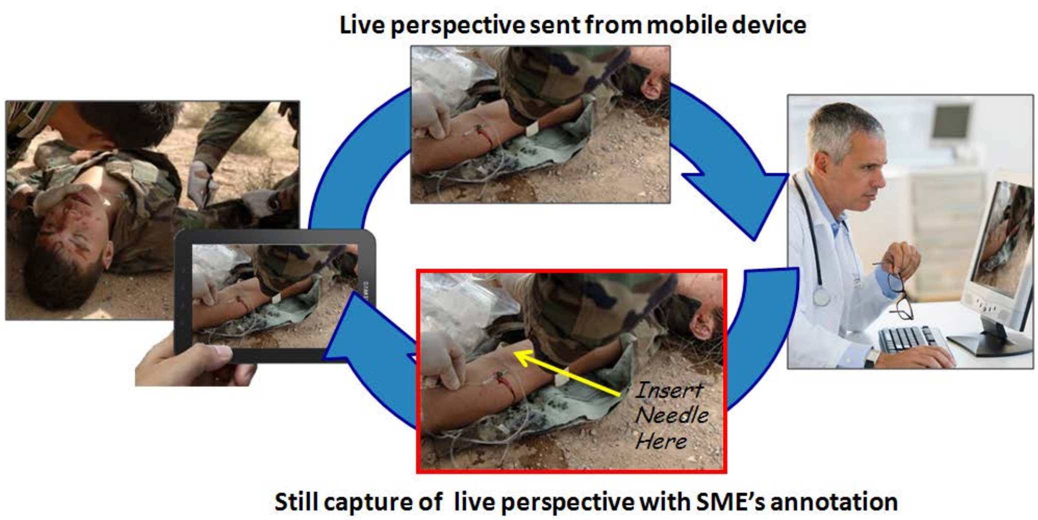


Figure 3.3: Perspective Sharing of Event and Receipt of Helper Markup

perspective data initially requires establishing a connection to a video capture device. Mobile devices are generally equipped with on-board cameras; however, depending on the circumstances, an off-board camera may be better suited for a collaborative task. Accordingly, the newly designed mobile application was implemented to accept a video capture device signal from either an embedded camera or an external camera. The external camera source can be either wired or wirelessly transmitting through TCP/IP. Recent development of small packaged video capture devices with integrated wireless transmitters, intended to be worn on the head, would be an ideal candidate for external connectivity. The degree of flexibility in video sources enables the mobile application to be scalable in order to address the various demands and in-field capabilities. The Worker can determine the video capture device source dynamically through the use of a camera selection interface. The camera selection interface presents to the Worker a graphical user interface choice of a "Remote Camera" versus an "Internal Camera", as shown in Figure 3.4.

Once a video capture device connection has been established, the active workspace is captured through frequent sampling, performed in a thread, of the camera's field of view. The captured contents are saved to an image buffer for processing. The sampling frequency can be determined

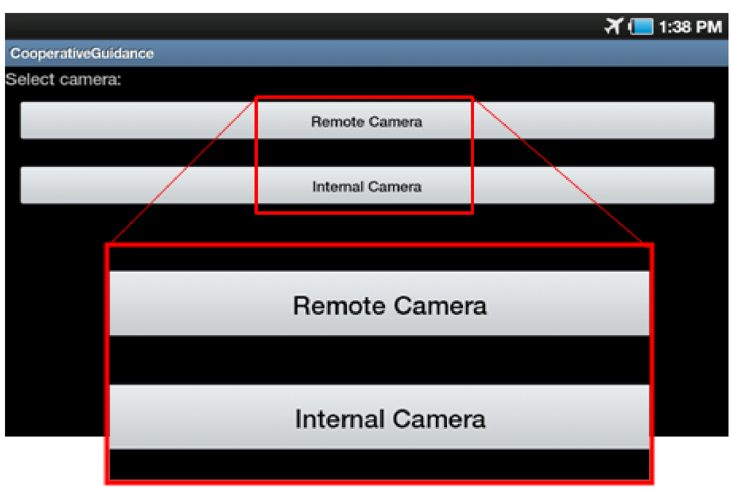


Figure 3.4: Video Capture Device Source Selection

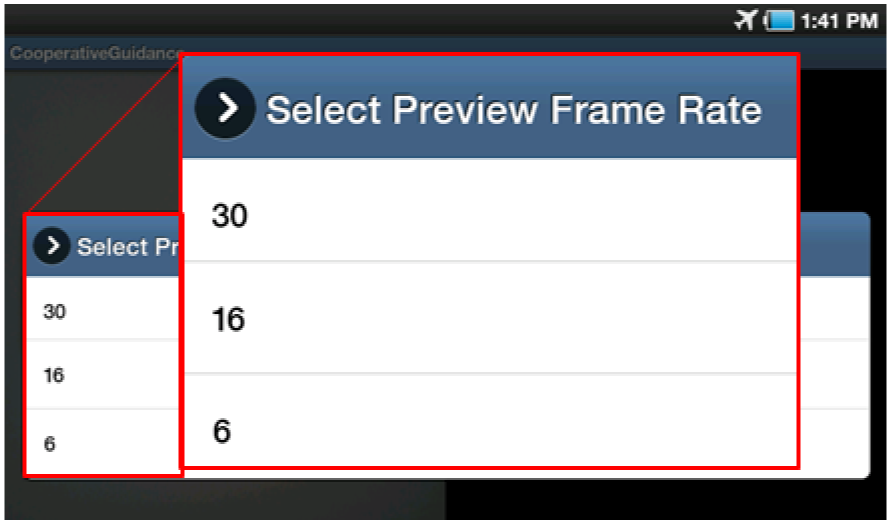


Figure 3.5: Video Capture Frame Rate Selection

by either the Worker and/or the Helper in real-time to adjust visual communication fidelity if power conservation is necessary. For example, if the battery was running low either collaborating member can reduce the visual frames per second shared, which could prolong the battery state of charge. The Worker is able to change the frequency through the Control Menu, as shown in Figure 3.5, and the Helper can change the frequency through the communication status TCP/IP message.

The configurable sampling and sequential transmission of the image is scalable to conserve power consumption from network utilization and resource processing. The camera's acquisition and image transmission can be selected at an upper limit of 30 frames per second and can be adjusted down to a lower limit of 6 frames per second. An additional power saving feature that the mobile application supports is variable image format conversion and compression. Prior to network transmission, the image buffer is processed to improve network utilization as well as to maximize the receiving parties' ability to handle the image without preprocessing the incoming data. The mobile application queries the mobile device to determine which image formats the device currently supports. Depending on which version of the Android Operating System is running on the mobile device, supported image formats may differ. Once the supported image formats are determined, the Control Menu's "*Select Preview Format*" is updated for user selection. A sample of formats supported includes (but is not limited to): JPEG, PNG, NV16, NV21, RGB565, and YUY2. The default image format that the image buffer is converted to is the Joint Photographic Experts Group (JPEG) format. The conversion utilizes the YuvImage() class which extends from the Android Graphics Object. The rectangular region of the display camera source is passed into the compression method along with a byte array output stream buffer to which the compressed data is written.

Following image conversion, the image data is encapsulated into a datagram package(s) for dissemination. The package(s) is transmitted through the mobile device's integrated network interface card (NIC) using standard Internet Protocol (IP) User Datagram Protocol (UDP). The Worker can alter the recipient(s) of the image through the Control Menu by modifying the destination IP address, as shown in Figure 3.6.

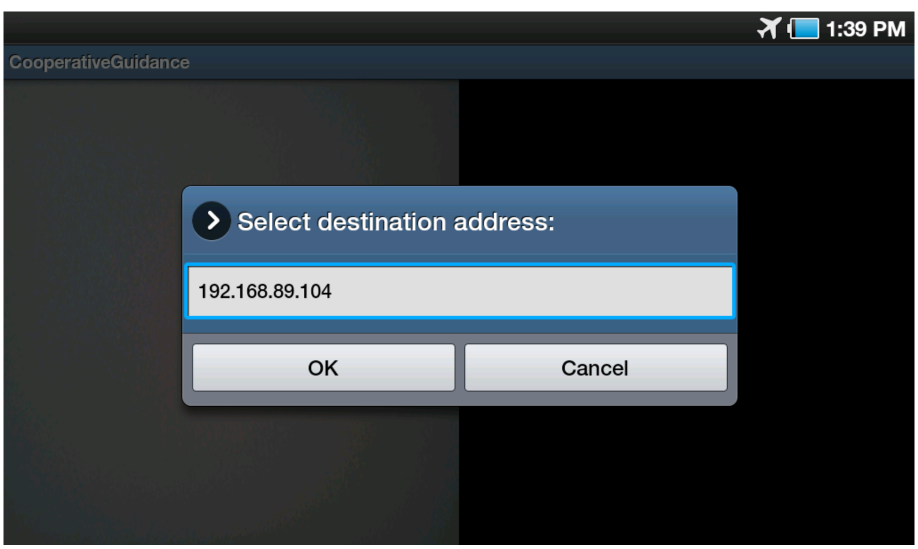


Figure 3.6: Remote Recipient of Worker's Workspace Images

Received images from the Helper are processed within a dedicated socket thread. Image data read from network datagram packets are stored into a byte array. The byte array is then decoded using the Android Graphic's BitmapFactory() class, which creates a bitmap from the byte array contents. The resultant bitmap is then forwarded to the appropriate image surface view on the mobile device's display for Worker's viewing.

The logical flow of the operations that *View Sharing* executes is displayed in the flow diagram below:

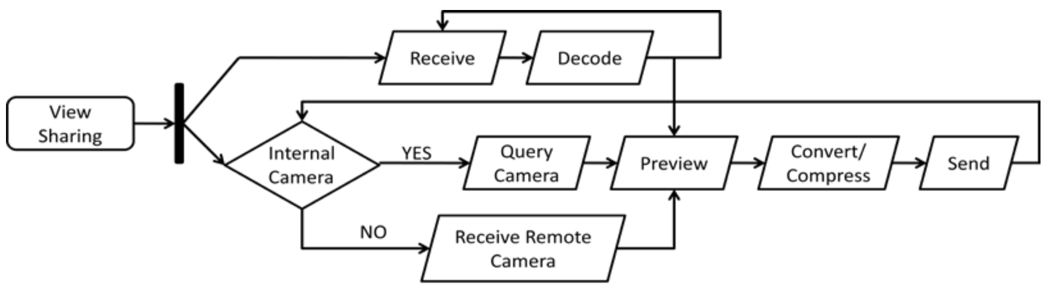


Figure 3.7: View Sharing Control Flow Diagram

3.3.2 Configurable Preview Modes

The mobile application utilizes multiple surface views to render previews of the live video data, regardless of source, and receives still image data from the remote Helper on the mobile device's display. This ensures that associated orientation and scene contents, appropriate for a given task, are being properly captured. Two overarching presentation modes were designed into the mobile application, full-screen and split-screen, to support versatile displays that could improve human effectiveness while executing cooperative tasks, as shown in Figure 3.8. During runtime, the Worker can change the preview mode to his/her preference through the Control Menu option "*Toggle Fullscreen*". In split-screen mode, the Worker sees the live video feedback on the left half of the screen and the Helper's annotated image on the right half of the screen. This configuration can serve as a reference perspective where the Worker may refer to the Helpers annotated image and apply the instructional information to the live adjacent perspective. In full-screen mode, the Worker sees the live video feed with a translucent overlay of the Helper's annotated image in the middle of the mobile device's screen, explained further in section 3.4.5, *Transparency Overlaid Preview*. This configuration can serve as a guide to the Worker as the Helper's annotation markups are merged with the live perspective.

Figure 3.8: Mobile Application Presentation Modes

The flow control for modifying the preview mode is shown in Figure 3.9.

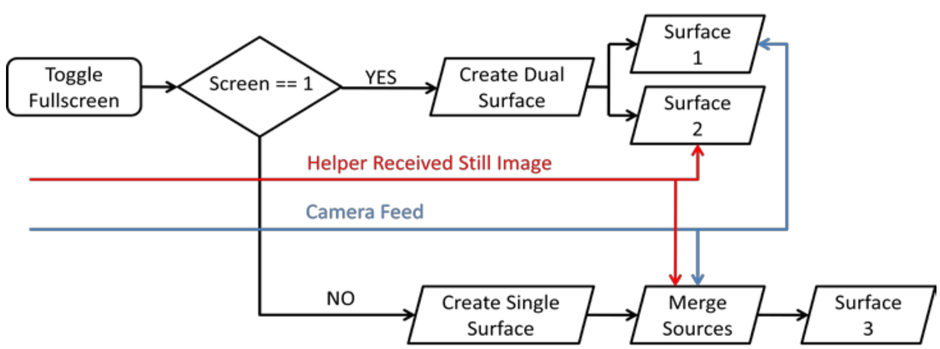


Figure 3.9: Configurable preview mode control diagram

3.4.2 Text Messaging

"Text Messaging" was implemented into the mobile application to support an alternative communication capability between Helper and Worker to relay guidance. Textual information received from remote Helpers is displayed on the mobile device's screen for the Worker's viewing. Interacting with active interface controls, namely the touch screen display and feedback sensors (e.g. speaker and tactile vibrators), textual information can be displayed and acknowledged between the cooperative pair(s). Received messages, exchanged using a threaded UDP socket, are read from the NIC into a local buffer. The contents of the buffer are then presented to the mobile display towards the upper section of the interface. To prevent overloading the Worker with rapidly changing messages, an acknowledgment message is implemented. The Worker acknowledges a message through a GUI button that generates an UDP ACK message to the remote Helper indicating the message was processed and the Worker is ready for a new message. A flowchart of the text messaging communication capability is displayed below in Figure 3.10:

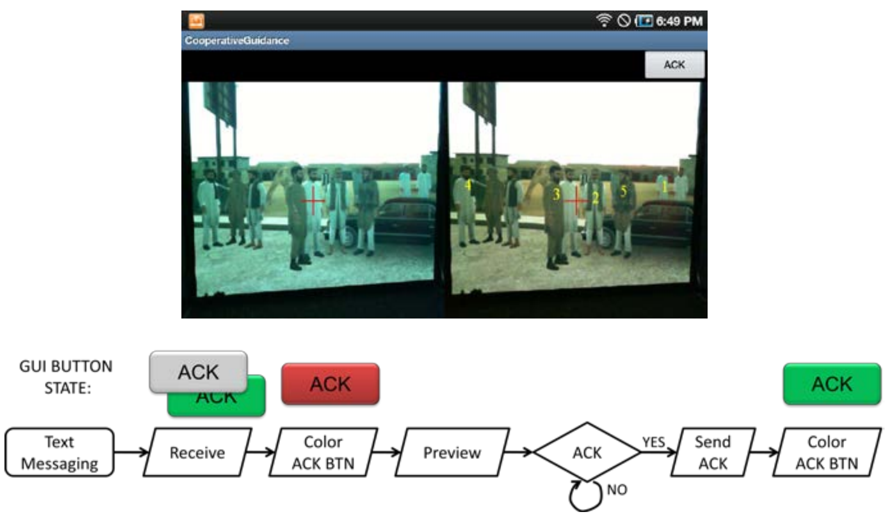


Figure 3.10: Text message control diagram

3.4.3 Audio Messaging

"Audio Messaging" supports the ability to transmit/receive audio information between Worker and Helper. This communication capability interfaces with both the mobile device's microphone and speaker hardware. The mobile application implements full-duplex audio communication across TCP/IP. The input and output audio signals are handled and are processed in separate

Figure 3.11: Guidance provided through audio messaging

threads to support simultaneous use if needed. For input, several features were built into the mobile application that allows the Worker to choose the most appropriate mode for capturing input audio sources. The first mode continuously captures “hot mic” input and transmits audio to the Helper. The second mode supports capturing and transmitting audio input only while depressing an external push-to-talk (PTT) button connected to the mobile device. Both input means were included in the design to address hands-free operations and power consumption considerations. In addition to the Worker initiated audio capturing and transmission, the mobile application permits external control of audio capturing and transmission through a TCP/IP socket trigger. This feature permits the Helper to enable/disable audio transmission remotely. The mobile application can receive stereo or mono inputs and support a wide range of frequencies (e.g. 11 KHz, 22 KHz, 44 KHz, etc.) and sampling rates (e.g. 8 bits/s, 16 bits/s, etc.) of audio sources to accommodate the numerous military and/or commercial headsets that may be connected and utilized with the mobile device.

Input audio capturing is done through a persistent thread which monitors the Worker’s ambient environment through the mobile device’s embedded microphone. Alternatively, if a headset with integrated microphone is connected to the microphone/headset jack, then its external microphone is used. The microphone’s captured audio signals are sampled at a configurable rate and frequency. The default setting for recording is 16 bits per sample at a sampling rate frequency of 11 KHz. Captured audio data are saved to an audio track memory buffer and passed to a pulse code modulation compression method prior to network transmission. The resulting audio data are written into a datagram package and transmitted through the mobile device's NIC to the predetermined host as identified through the Control Menu.

Output audio playback is performed through a separate thread that monitors remote audio communication coming into the mobile device via the NIC. The thread will receive network traffic and place the information into a memory buffer. The contents of the network data are processed through an uncompressing method and the resulting data are written to an audio track memory buffer. The audio track data are a playable audio format and are sent to the mobile device's audio interface for rendering. If the audio interface is currently in use, the new audio track is queued until the audio interface is able to perform its playback. The playback will occur on the mobile device's internal speaker in the absence of a connected headset; otherwise the playback will occur in the connected headset.

An audio control flowchart diagram is displayed in Figure 3.12:

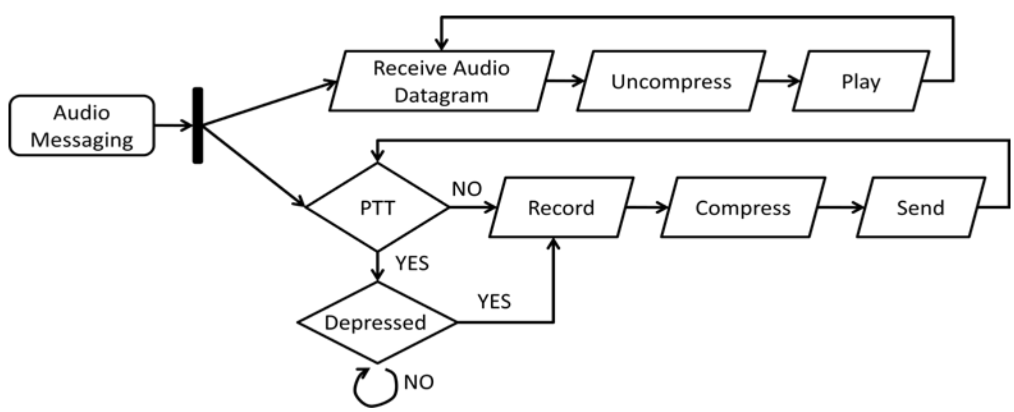


Figure 3.12: Audio messaging control diagram

3.4.4 Dynamic Image Alignment

"Dynamic Image Alignment" enables the alignment of received images from a Helper to the Worker's active workspace's orientation displayed on the mobile device. Mobile devices' form factors afford them to be portable and perform on-the-move processing. However, their compact size can make them difficult to hold static as users manipulate and interact with the device. When collaborative information is captured in markup images, offsets between the current live perspective and the captured markup perspective may slightly differ. In the current design, when merging the live and markup images in the full-screen mode, a ghost effect could be rendered if the two orientations do not align, as is depicted in Figure 3.13.

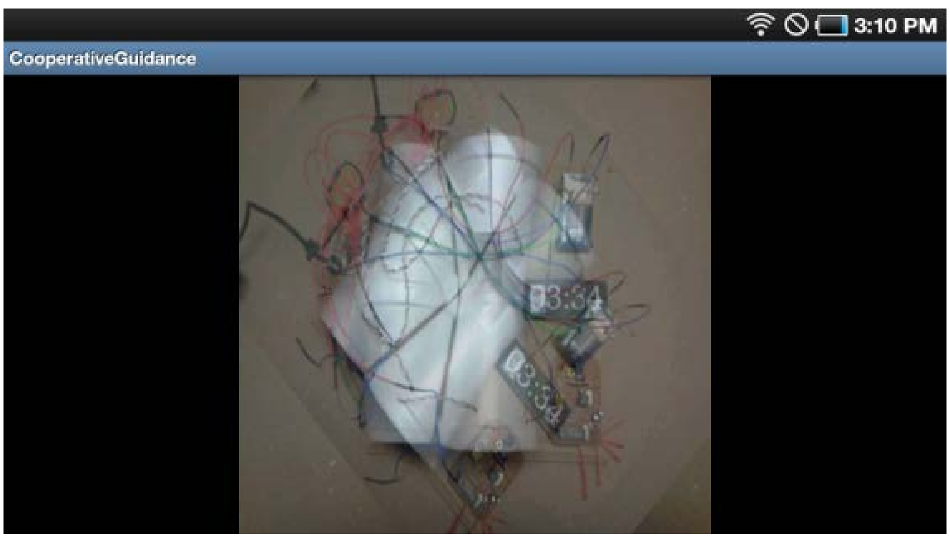


Figure 3.13: Image alignment ghosting effect

To prevent this visually distracting effect, the ability to align the Helper's still markup images with the active live perspective is desirable. For this collaborative Android system, alignment was implemented in power conscious software and hardware approaches. There are numerous computer vision software techniques that can perform feature extraction and image transformation allowing the annotated markup image to align/register to the live perspective captured on the mobile device. However, for this implementation, power conservation and user performance abilities were driving factors. Therefore, the software method implemented leverages the Worker's physical input from the touch screen to manipulate the rotational alignment of the Helper's still image to the desired orientation as the Worker needs it. Several factors contributed to this design approach. First, if the alignment process was automated using traditional computer vision registering operations without the user's initiation, then the overall alignment capability would consume resources regardless of Worker's need. For example if the Worker is moving, working on the physical task, or not focusing on the mobile device displays, then utilizing the mobile device's resources to perform image alignment is not ideal for power savings. Second, if the Worker is focusing on the mobile device's display and attempting to comprehend the graphical information from the Helper's annotated image, there is a high probability that the mobile device will not remain stationary. This movement, albeit nominal to the human, can produce jitter effects as the image registration processes attempt to improve alignment throughout the movement. This jitter could cause additional unnecessary workload on the Worker. The inclusion of filters or conditional preprocessing prior to invoking the image registration process would improve or address these jitter artifacts. However, they still require the use of resources to calculate the filter and conditionals repeatedly.

For these reasons the software approach establishes an on-touch callback process that is activated when the user is actively touching the mobile device's screen. From the Worker's touch placement and movement on the screen the callback process interprets them to update an angle of rotation *degree* variable. The Worker's touch inputs are relative inputs, meaning the Worker can continue an angular rotation through several finger movements in the same direction on the mobile devices screen that do not need to be continuous. These gesture inputs add or subtract from the current displayed image orientation. Upward touch movements subtract degrees and downward touch movements add degrees to the current angle of rotation, as illustrated in Figure 3.14. The updated *degree* variable is then processed through the Android Bitmap class where the orientation to the still image is applied.

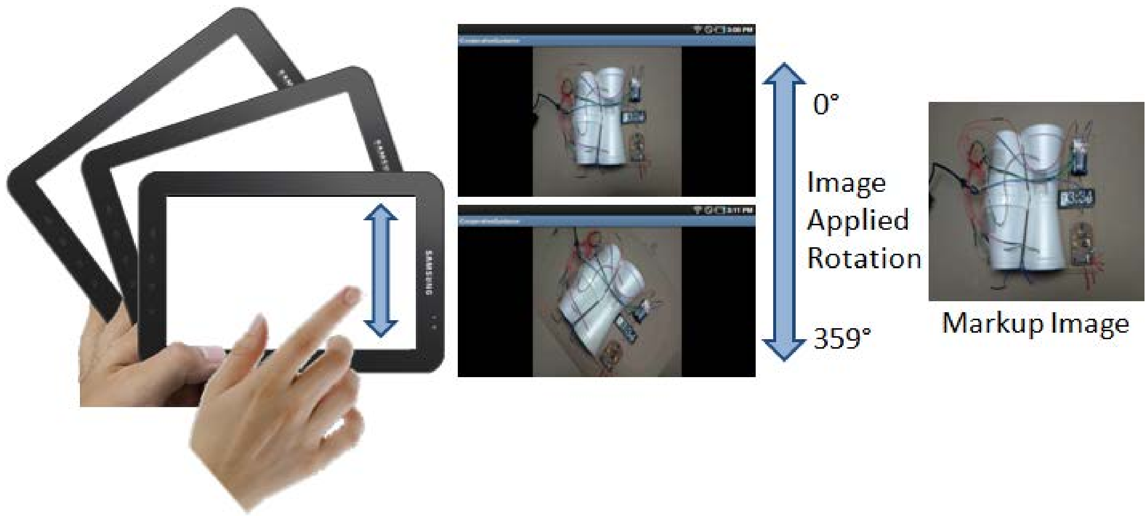


Figure 3.14: Alignment of Helpers still markup with active view on mobile device

The hardware approach queries the integrated gyros on the mobile device to determine the device's current orientation. In the view sharing section, the orientation of the mobile device, as assessed from the internal gyro sensors, is transmitted along with the camera's generated images. Once an image is received from a Helper the image's corresponding angle is referenced and an offset is calculated. The calculated offset angle is applied to the Helper's image and displayed to the Worker.

In addition to the onboard calculated image alignment methods, image registration can be offloaded to the Helper. On the Helper's side, angular rotation can be determined and then the oriented image can be transmitted back to the Worker.

A control flowchart illustrating the image alignment process is displayed below in Figure 3.15:

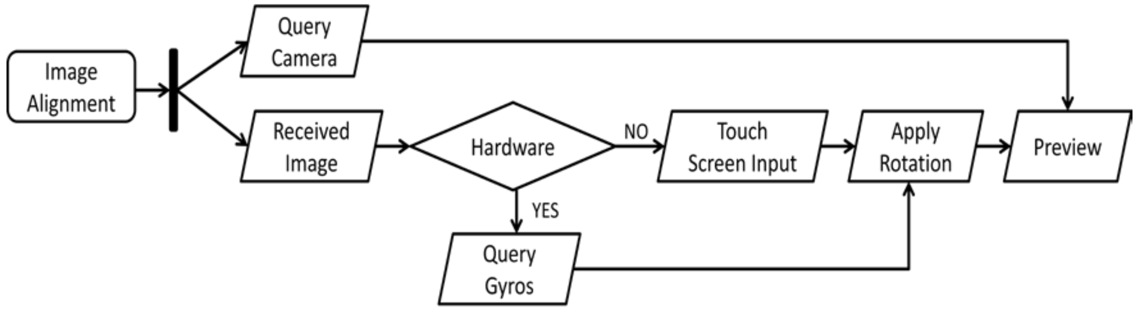


Figure 3.15: Image alignment control diagram

3.4.5 Transparency Overlaid Preview

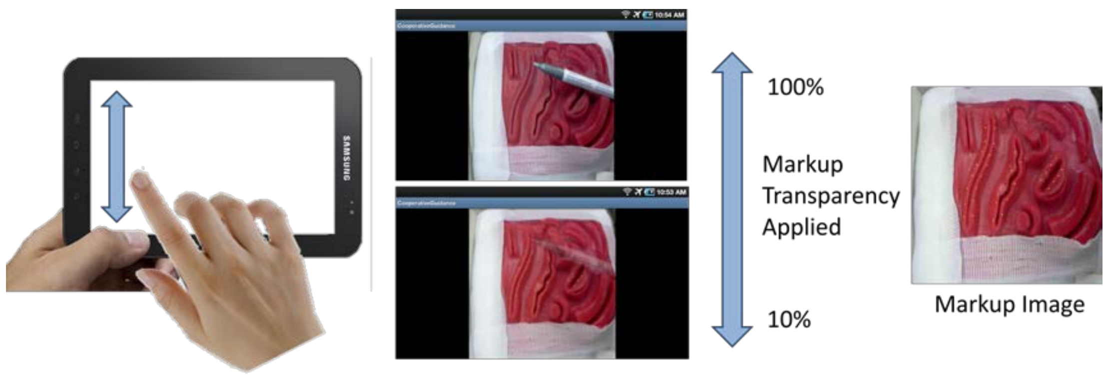


Figure 3.16: User configurable transparency of markup image

“Transparency Overlaid Preview” is an additional capability that extends the View Sharing application. The Worker has the option to dynamically adjust the transparency of received images from a Helper, which are overlaid on top of the live perspective captured by the camera in a separate view surface. A network thread monitors incoming images from the Helper. Upon receipt, image data is saved to a bitmap by processing the received UDP datagram(s) data through the Android Graphic’s BitmapFactory() class. The resulting bitmap is then sent to the overlay view surface to display the received image to the Worker. To adjust the image’s transparency, a touch screen callback process is used. The use of the touch screen is designed as the input source, instead of dropdown menus, GUI buttons, and/or keyboard inputs, for ease use as well as to minimize the cognitive burden to the Worker. The Worker can adjust the transparency of the overlay by using pan gestures on the left half of the overlay surface view that resides above of the camera feed. Panning up reduces the transparency, and panning down increases the transparency, as shown in Figure 3.16. With these gestures, the Worker can quickly set the transparency of the overlay to a level suitable for the current task. Note that the implementation chosen to control the overlay's transparency levels do not occupy any space on the user interface and thus do not distract or clutter the mobile device's display of the active workspace. A control flowchart is displayed below in Figure 3.17:

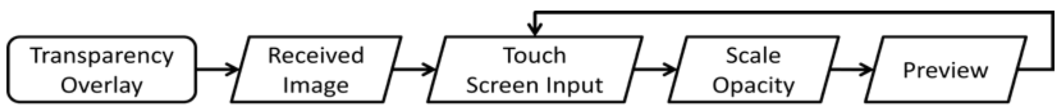
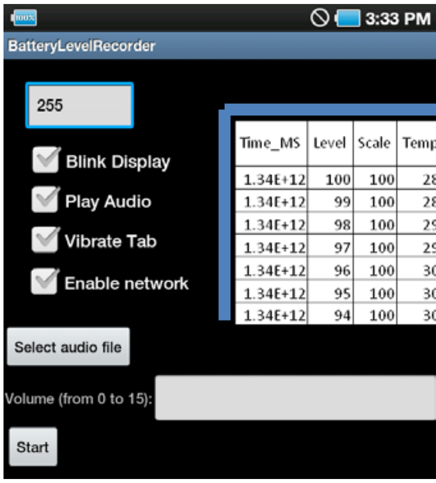


Figure 3.17: Transparency overlaid preview control diagram

3.5 Assessment of Modalities in Power Consumption

To investigate the effects that the implemented communication capabilities have on the mobile device, an evaluation algorithm was developed to measure the power consumption and time duration under isolated modality loads. The evaluation algorithm was designed to interface with the Android power manager, which initiates a notification whenever the battery state changes. Additionally, the algorithm monitors processor usage by querying the processor during runtime and records various device parameters in order to construct power profiles. To quantify the power effects of each implemented multi-modal capability as described above, the power profiling algorithm was designed to run as a background process so that it could be run in conjunction with other mobile applications or by itself to obtain power consumption data. The main objective in the development of this application was to help users (Helpers and Workers) manage power consumption to support remote collaboration. Specifically, the goal was to prevent premature expiration of power prior to the completion of the collaborative task. This was achieved by enabling both Helpers and Workers to monitor the Worker's mobile device's power state and evaluate the impact of operating communication capabilities on the power consumption rate and consequent remaining duration of device operation.

Effective management of power consumption is paramount for distributed cooperative tasks that require coordination between Workers and Helpers. Critical to the communication link between the pairs is the mobile device's finite power that enables the various multimodal capabilities needed in the dissemination and receipt of task procedures and pertinent data. By measuring and classifying the usage penalty per capability with their respective resolution and fidelity settings, the operating time can be calculated, allowing the development of a power measurement process (PMP). The PMP was implemented to isolate and assess the energy cost of the mobile device's features that support collaborative communication capabilities. For example, the communication capability *View Sharing* leverages the mobile device's display, NIC, and camera. In order to assess the total power consumption for *View Sharing,* data collection on the power consumed by each of the sub-features was performed in isolation and then their cumulative power effect was assessed. The PMP isolates the mobile device features through the use of a simple graphical interface. The interface enables the features and assigns fidelity settings for each, such as the display brightness value and refresh rate. Upon configuring the features, power measurements are initiated through the start button on the interface, and data are collected until the mobile device shuts down due to running out of power. The collection of data is triggered through an *Intent* object that monitors the power manager services for the ACTION.BATTERY.CHANGED flag. When the flag is set, the battery's current state of charge has changed. Accordingly, the PMP logs several mobile device values to construct a power profile for the feature under evaluation. The values recorded are: time since start, battery level, temperature, voltage, CPU_USER, CPU_NICE, CPU_SYS, CPU_IDLE, CPU_IOWAIT, CPU_IRQ, CPU_SOFTIRQ, CTXT, and number of processes running. Figure 3.18 shows the PMP interface along with an example log profile.



Time_MS	Level	Scale	Temp.	Voltage	CPU USER	CPU NICE	CPU SYS	CPU IDLE	CPU IOWAIT	CPU IRQ	CPU SOFTIRQ	CTXT	Procs Running
1.34E+12	100	100	280	4164	3438	42	3139	70659	197	26	53	347031	3
1.34E+12	99	100	289	4094	3696	42	3294	71659	199	29	53	358925	8
1.34E+12	98	100	294	4085	8465	42	9446	91791	199	137	53	578600	7
1.34E+12	97	100	297	4074	13455	42	15787	112307	200	250	54	803310	7
1.34E+12	96	100	300	4058	18485	42	22184	132727	200	375	54	1029061	8
1.34E+12	95	100	303	4048	23464	42	28465	153115	201	496	54	1253980	5
1.34E+12	94	100	306	4050	28374	42	34725	173589	201	621	54	1477275	7

Figure 3.18: Power measuring process interface and example log output

The objective of enabling Helpers to remotely monitor power levels and negotiate with Workers on appropriate communication capabilities is facilitated by the status message. The status message, integrated into the mobile application's implementation, shares the current battery state of charge and all active communication capabilities with associating fidelity levels between Worker and Helper. The current state of charge and the ability to calculate the cumulative power draw that the active communication capabilities are using enables the Helper and Worker to predict the expiration time of the battery. Informed of the mobile device's power condition, either party can suggest appropriate feature level changes, if needed, to ensure the battery is not prematurely depleted. Capabilities and features of the mobile device can likewise be adjusted real-time through the status message. This distributed control enables either the Worker or Helper to set the mobile device to a power saving mode. If the Worker modifies any of the device's settings, those changes are performed without acknowledgment and occur immediately; implemented changes are then reported to the Helper. When triggered by the Helper, the mobile application requires the Worker to acknowledge the suggested changes in feature settings before they are applied. As environment and context surrounding the mobile user may limit the effectiveness of the remote guidance (due to, e.g., bright lights, loud noises, etc.), it is reasonable to give the Worker the ultimate choice and the ability to actively negotiate modality changes, affording them a flexible approach to power conservation. A control flowchart is displayed below in Figure 3.19:

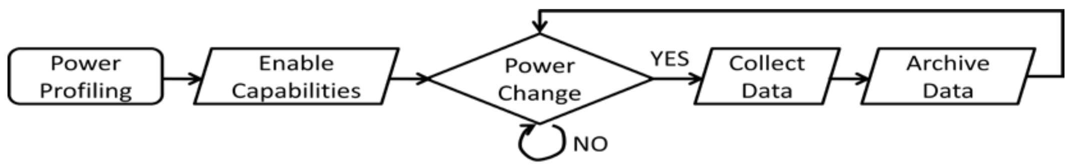


Figure 3.19: Power measuring process control diagram

4. RESULTS

The following chapter highlights the performance of the mobile application and its impacts on the individuals engaged in pair collaboration. The description of five distinct tests and the analysis of their results are reported, further discussion is provided in the next chapter. The practical examples presented are important to the United States Air Force and highlight the effects on user performance, confidence, and accuracy when collaborating with a distributed Helper to execute these specific tasks. The impact of the mobile application is analyzed in the following scenarios: improvised explosive device disposal, finding-fixing-tagging targets of interest, complex building block assembly, and medical treatment situations.

4.1 Improvised Explosive Device Defusing

A pilot demonstration involving the defusing of a simulated improvised explosive device (IED) was conducted to assess the extent to which the Android application supported remote collaboration. This task was selected because of its high relevance to current military operations and because IEDs are not standard in their design, having numerous wire configurations and trigger features. In short, defusing IEDs involve systemic sequential wire identification and disarming (cutting or rerouting wires) to make the IED inert.

4.1.1 Participants

Twelve participants volunteered for this study, 8 men and 4 women, ranging in age from 23-30 (M = 25) years. All participants had normal hearing and normal or corrected-to-normal vision.

4.1.2 Experiment Design

A within-subject design was employed with four levels of Modality Interface (Audio, Video with Markup, Video with Audio, and Video with Markup and Audio). The order in which each Worker utilized a modality was controlled by counterbalancing the usage order so as not to bias the experimental conditions. All Workers took part in a training session to familiarize themselves with the task and devices. The Workers trained defusing four IEDs per experimental condition. Workers were given the option for more practice trials; however, none of them felt the need for additional practice. The four experimental conditions and IED configurations were randomized per Worker.

4.1.3 Apparatus

Four simulated IEDs were used in the experiment. Each IED consisted of a clock, power source, control chip, and explosive charge containers as seen in Figure 4.1. There were nine wires on

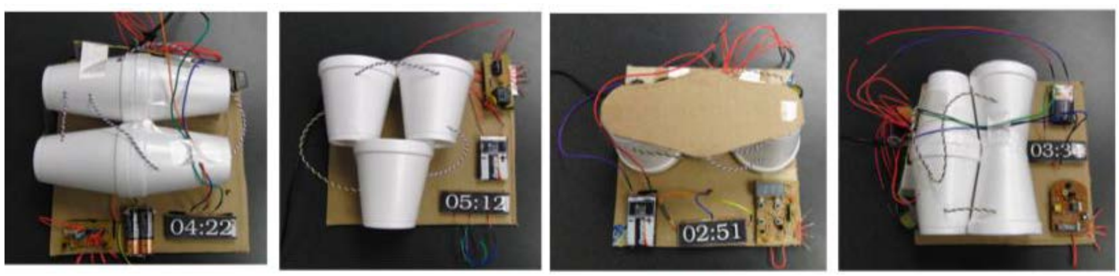


Figure 4.1: Simulated Improvised Explosive Devices

each IED: seven were active and two were distracters. The Worker collaborated cooperatively with a remote confederate Helper who had detailed instructions for disarming each IED and experience communicating through the various multimodal communication capabilities. Workers used a Samsung Galaxy Tablet running the developmental Android application to interact with the remote Helper through a Wi-Fi connection. The Galaxy Tablet was mounted on a stand to allow the Worker to freely use their hands, as seen in Figure 4.2. The Helper was situated in front of a workstation which was isolated from the experimental area. The Helper's workstation allowed them to communicate via TCP/IP, capture, and annotate images from the Worker's tablet to assist them in their task.

4.1.4 Procedure

Four conditions were evaluated: 1) Audio only where the Helper could not see the Worker's workspace; 2) Video with Markup where the Helper monitored the Worker's workspace and provided markup directives; 3) Video with Audio where the Helper monitored the Worker's workspace and provided verbal directives; and 4) Video with Markup and Audio where the Helper monitored the Worker's workspace and could provide directives through both markup and verbal interactions.

In the Audio condition, Workers spoke to the Helper via VoIP where they had to describe the IED in order for the Helper to relay the proper sequence for disconnecting the active wires. The Video with Markup condition consisted of the Helper capturing a picture of the IED from the tablet's perspective, then annotating the picture in real-time on their workstation. The annotated image, which showed the order of wires to disconnect, was sent to the Worker participant to defuse the IED. The Video with Audio condition consisted of the Helper monitoring the Worker's perspective while supplying verbal instructions to defuse the IED. The Video with Markup and Audio condition combined the Audio and Video conditions so that the Helper and Worker were able to talk to each other as well as send annotated images.

For each condition, Workers defused a unique IED. They were asked to complete the task as fast as possible without making any errors. A countdown clock was used to impose time pressure, initially starting at one minute and decrementing each second.

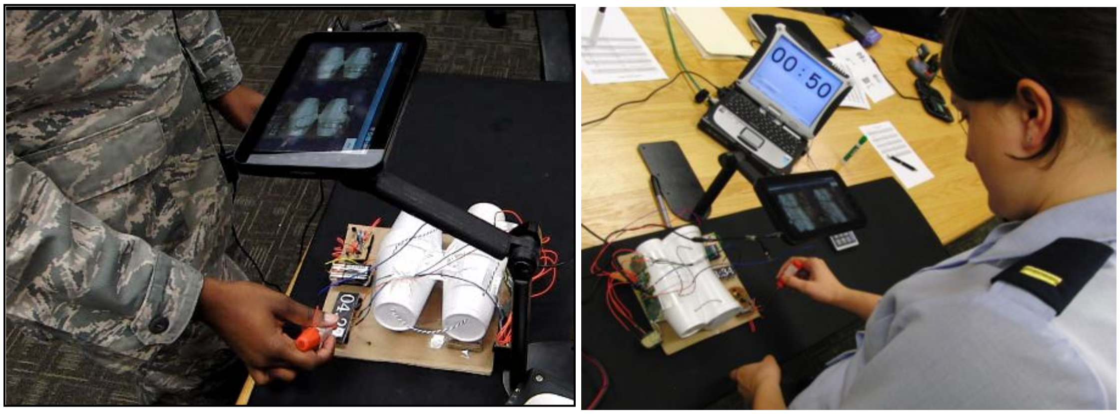


Figure 4.2: Participant diffusing IED with Tablet

4.1.4 Results

Mean task completion time and their respective standard errors for the four experimental conditions are displayed in Figure 4.3. A four condition repeated measures Analysis of Variance (ANOVA) of these data revealed a statistically significant main effect for conditions, $F(3,33) = 70.88$, $p < .05$. Subsequent post hoc Tukey-tests, with alpha set at .05, revealed that Workers using "Video with Markup" and "Video with Markup and Audio" completed the task statistically faster than in the other two modes, but were not different from each other. The Tukey-test also found that participants using Video with Audio were faster than Audio alone.

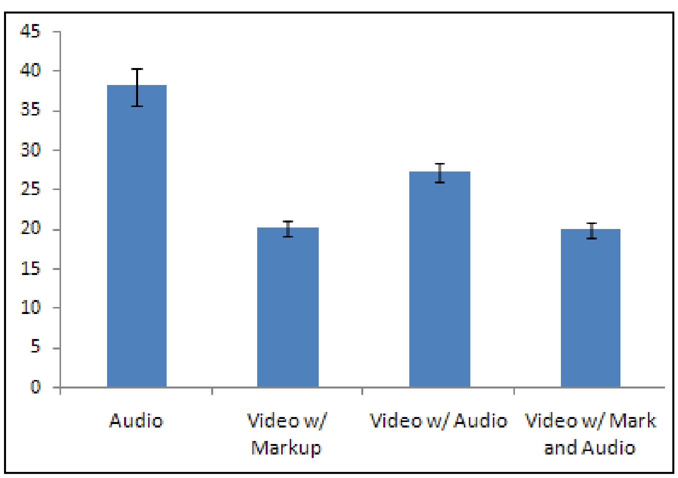


Figure 4.3: Mean completion times for each of the four experimental conditions.
Error bars are standard errors.

4.2 Find, Fix, and Tag Experiment and Results

An initial demonstration involving a find, fix, and tag task was conducted to assess the extent to which the developed Android application supported remote collaboration. Participants communicated with a remote expert using various modalities to complete the evaluation task. Task components involved: 1) identification of a specific individual in a crowd of people, 2) alignment of an aiming device on an identified individual, and 3) initiation of a tagging sequence. The modality interfaces investigated were Audio, Video with Markup, Video with Audio, and Video with Markup and Audio.

4.2.1 Participants

Eight military and four civilian participants volunteered for this study (eight men and four women) ranging in age from 23-30 (*M* =25) years. All participants had normal hearing and normal or corrected-to-normal vision. Additionally, all participants had prior training and experience in the usage and handling of a rifle. The participants collaborated with a remote Helper who knew the order and identity of the individuals being tagged.

4.2.2 Experiment Design

A within-subject design that was balanced using a Latin-square procedure was employed with four levels of Modality Interface (Audio, Video with Markup, Video with Audio, and Video with Markup and Audio). All participants took part in a training session to familiarize themselves with the task and devices. The Workers trained by communicating with the remote Helper and marking targets of interest with an AirSoft M-4 rifle per experimental condition. Workers were given the option for more practice trials; however, none of them felt the need for more. The four experimental conditions and virtual target configurations were randomized per Worker.

4.2.3 Apparatus

Each Worker used an affixed pivoting AirSoft M-4 Rifle with a camera attached to the forward barrel as shown in Figure 4.4.

Figure 4.4: Rifle with attached camera

Workers were instructed to stay behind a partition wall, which blocked their line of sight to the active scene, and utilize the rifle mounted camera's perspective for the task, as seen in Figure 7. The partition wall was positioned in front of an 8'x10' projection screen that rendered a virtual scene consisting of a gathering of 12 potential targets of interest.

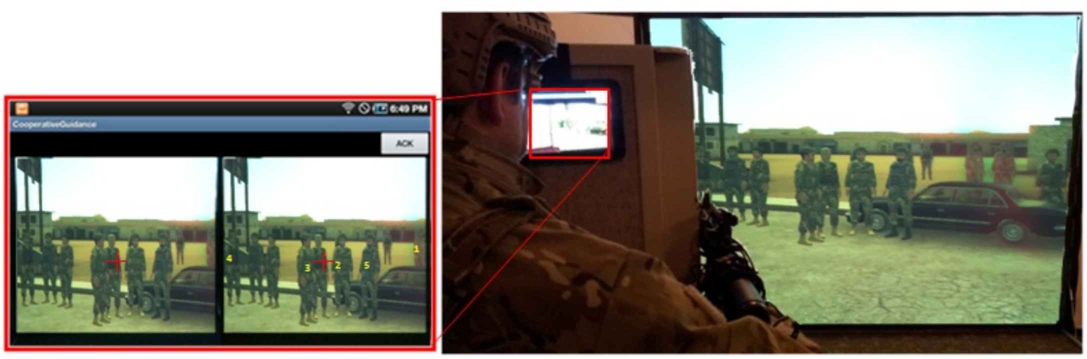


Figure 4.5: Structure and experiment scene

The rifle/camera provided a live video feed to a Samsung Galaxy Tablet running the developmental Android application. The Tablet was stationary mounted to the partition wall allowing the participants to freely use their hands, as seen in Figure 4.5.

The remote Helper communicated with the Worker through the tablet running the collaborative Android application through a Wi-Fi connection. They were situated a workstation, which was isolated from the experimental area, as shown in Figure 4.6. The Helper workstation allowed the cooperative pairs to communicate via streaming audio as well as capture and annotate still images from the Worker's tablet. The Helper used this tool to direct the Worker in finding and tagging the hostiles in a specific order.

Figure 4.6: Helper collaborative workstation

4.2.4 Procedure

The four conditions that were evaluated included: Audio only; Video with Markup; Video with Audio, Video with Markup and Audio. In the Audio condition, the Helper had to verbally describe to the Worker the characteristics of the individual that required tagging. The Helper's description of the individual started with a clothing description, an indication of facial hair, and whether the individual was wearing anything on their head. The Video with Markup condition consisted of the Helper capturing a picture of the participant's perspective from the rifle mounted camera then annotating the picture in real-time on their workstation. The annotated image, which showed the order of individuals to be tagged, was sent to the Worker to initiate the tagging action. The Video with Audio condition consisted of the Helper monitoring the participant's perspective while supplying verbal instructions regarding the individual to be tagged. The Video with Markup and Audio mode combined the Audio and Video conditions so that the Helper and Worker were able to talk to each other as well as send annotated images.

For each condition, participants tagged unique individuals. They were asked to complete the task as fast as possible without making any errors.

4.2.5 Results

Mean task completion time and standard errors for the four experimental conditions are displayed in Figure 4.7.

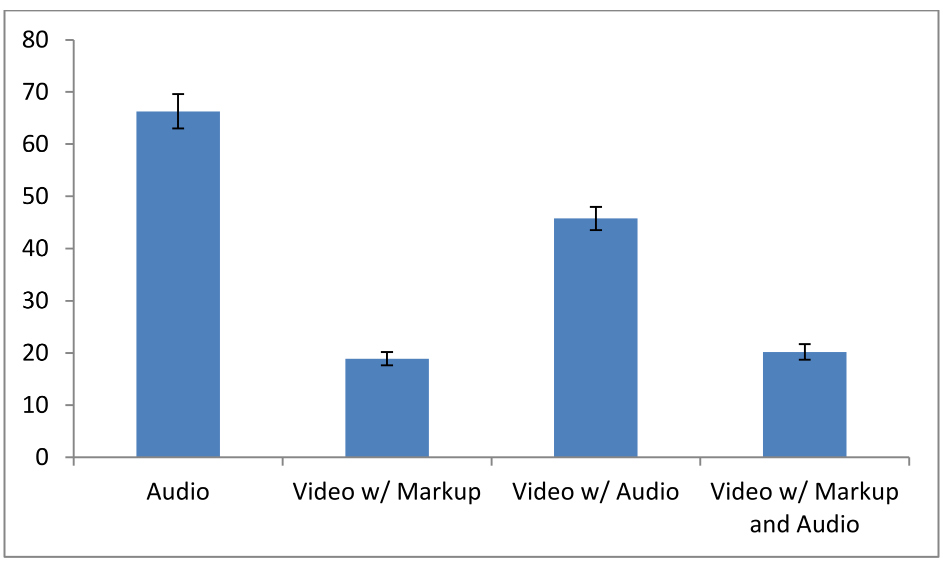


Figure 4.7: Mean completion times (sec) for each of the four experimental conditions. Error bars are standard errors

A four condition repeated measures ANOVA of these data revealed a statistically significant main effect for conditions, $F\ (3,33) = 70.41$, $p < .05$. A subsequent post hoc Tukey-test with alpha set at .05 revealed that participants using Video with Markup and Video with Markup and Audio completed the task statistically faster than the other conditions, but were not different from each other. The Tukey-test also found that participants using Video and Audio were statistically faster than Audio alone.

Mean accuracy and standard errors for the four experimental conditions are displayed in Figure 4.8.

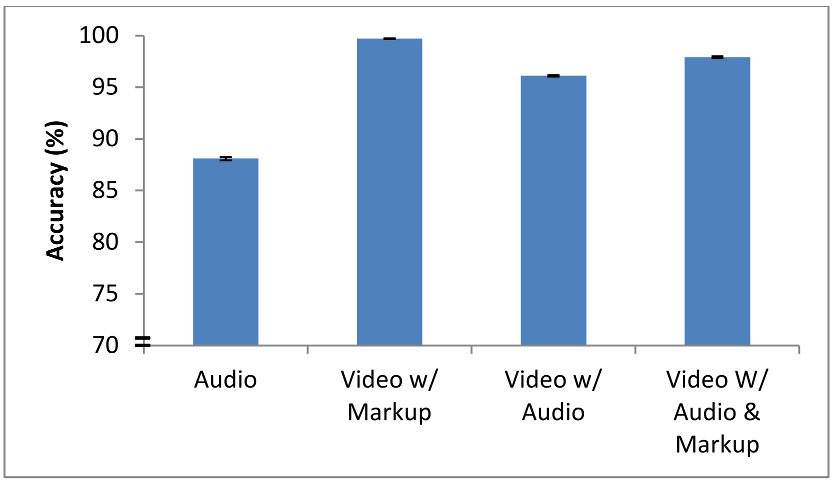


Figure 4.8: Mean accuracy for each of the four experimental conditions. Error bars are standard errors

A four condition repeated measures ANOVA was performed on these data and revealed that the mean accuracy values in the four conditions did not statistically differ from each other, F (3,33) = 2.24, p > .05. Additionally, the degree to which the experimental conditions affected the total verbal communication time was evaluated. It was found that the style and amount of verbal information relayed between cooperative pairs differed when a shared visual perspective was available. Figure 4.9 shows the mean voice usage times the remote Helper required to achieve common ground in positively identifying the experimental targets. A t-test revealed that the

Audio condition required more communication time then the Video w/ Audio, t (7) = 4.27, $p < .05$.

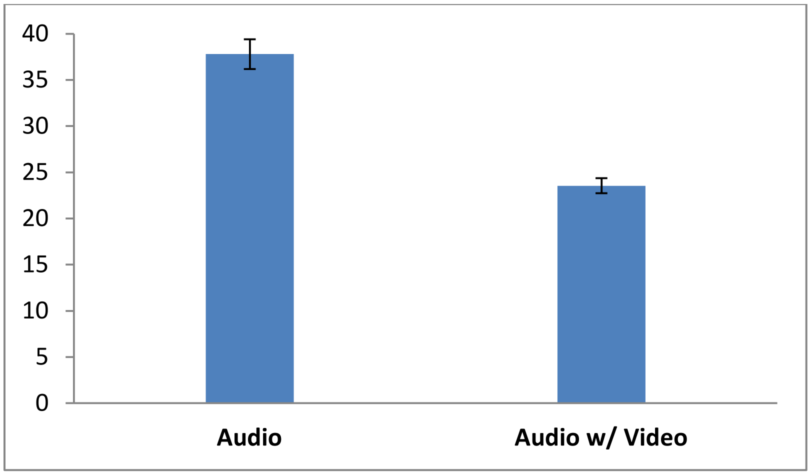


Figure 4.9: Comparison of active voice usage time (sec) of audio conditions

4.3 Building Block Team Assembly

An evaluation of team performance and their ability to effectively communicate while constructing a multi-level abstract structure from building blocks using the mobile application was performed. This task was selected because of its high degree of negotiating between Worker and Helper cooperating towards an end goal. This type of task requires detailed collaboration for block identification, orientation alignment, and location placement.

4.3.1 Participants

Volunteers for this study included 32 participants (17 men and 15 women) ranging in age from 23-30 (M=25) years. The participants teamed up in pairs of two, consisting of a Worker and a Helper, collaborating using various modalities to complete the building task. All participants had normal hearing and normal or corrected-to-normal vision.

4.3.2 Experiment Design

A within-subject design that was balanced using a Latin-square procedure was employed with thefour levels of modality interface (Audio, Video with Markup, Video with Audio, and Video with Markup and Audio). All participants took part in a training session to familiarize themselves with the task and devices. The teams trained by collaboratively communicating with each other to construct practice models per experimental condition. Teams were given the option for more practice trials; however, none of them felt the need for more. The four experimental conditions and building model configurations were randomized per team.

4.3.3 Apparatus

Sixteen building block guides were used in the experiment. Each guide consisted of 46 pieces and had three levels. The model pieces illustrated in the guides were randomly selected from a total of 108 pieces that consisted of eight colors (orange, black, blue, red, yellow, brown, dark green, and lime green) and six sizes (1x2, 1x3, 1x4, 2x2, 2x3, and 2x4 studs) The teams worked cooperatively to identify and place blocks onto a green board that measured 10 inches by 10 inches. Building blocks were located in a pile next to the green board approximately 5-8 inches to the right. Worker used a Samsung Galaxy Tablet running our developmental Android application to interact with the Helper through a Wi-Fi connection. The Galaxy Tablet was mounted on a stand above the green board to allow the participant to freely use their hands, as seen in Figure 4.10.

Figure 4.10: Worker’s Mobile Device Apparatus

The Helper was situated in front of a workstation, which was isolated from the experimental area. The Helper’s workstation allowed him/her to communicate via TCP/IP, capture, and annotate images from the Worker’s tablet to assist them in their task. The Helper’s annotations consisted of free form shapes that were filled with selectable colors, as shown in Figure 4.11.

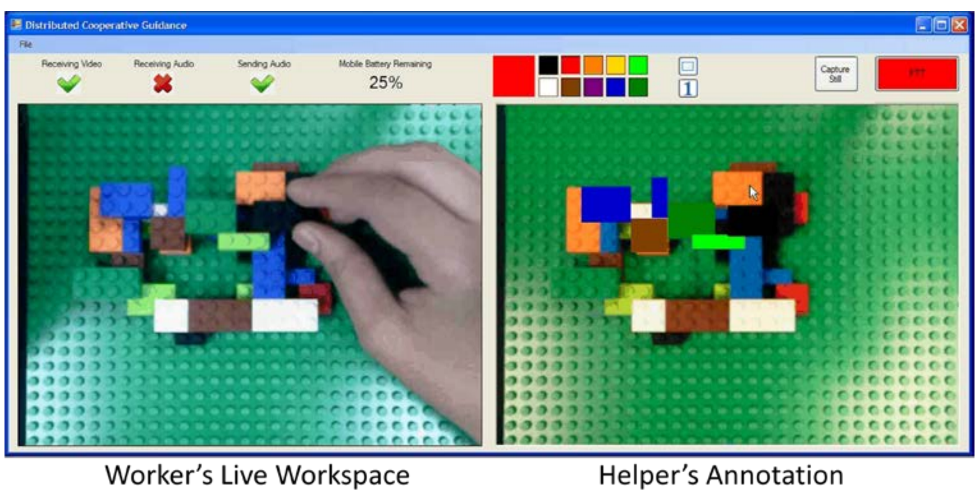


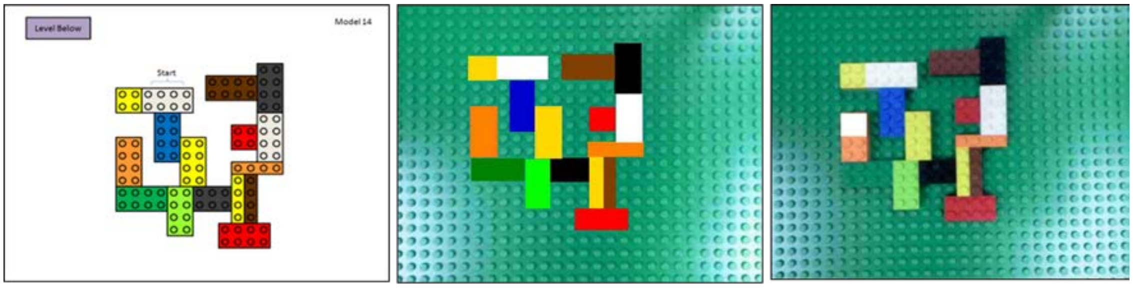
Figure 4.12: Helper's Workstation

4.3.4 Procedure

Teams of two, consisting of a Worker and a Helper, collaborated using various communication modalities to complete the building task. The modality interfaces investigated were Audio, Video with Markup, Video with Audio, and Video with Markup and Audio.

In the Audio mode, the Helper had to verbally describe the color, size, orientation, and placement of the building blocks to the Worker from the active build guide, shown in Figure 4.12 (a). The Helper's instructional dialogue describing the block and placement was not restricted in any manner, and it was left up to the teams to generate their unique shared common language used in the building process. The Video with Markup condition consisted of the Helper capturing a still picture of the Worker's live perspective from the mobile device's integrated camera. The still image could then be annotated in real-time on the Helper's workstation. The annotation process required the Helper to select the color used in the annotation, followed by clicking and holding the left mouse button down while dragging until the desired shape was illustrated. Upon releasing the left mouse button, the markup annotation was fused with the still image and transmitted to the Worker, as shown in Figure 4.12(b). The Helper could undo their annotation by selecting the right mouse button. The undo process could be applied five times to clear past annotations. If five corrections were not sufficient, the Helper could recapture a still image and apply fresh annotations. The Video with Audio condition consisted of the Helper monitoring the Worker's perspective while supplying verbal guidance to describe and place the current building block properly in the model. The Video with Markup and Audio condition combined the Audio and Video conditions so that the Helper and Worker were able to talk to each other as well as send annotated images.

For each condition the team members were asked to complete the task as fast as possible without making any errors. Examples of completed tasks are shown below.



(a.) Building Guide (b.) Helper's Annotation (c.) Worker's Workspace

Figure 4.12: Reference Guide, Helper's guidance to Worker, and Worker's execution of guidance

4.3.5 Results

Team performance was analyzed with regards to completion time, while the effectiveness of the collaborative tool was measured by the amount of data transmitted between team members to complete the task. Perceived mental workload was also collected using the NASA-TLX. All teams achieved accuracy of the building task of at least 97.5 % while completion time was used to assess team performance. A statistically significant main effect was found for completion time across the four experimental conditions, F (3, 42) = 34.2, $p < .01$. Post hoc test found that teams completed the building task significantly faster in the Video with Markup and Audio (M = 625.0 sec) condition as compared to Video with Markup (M = 735.1 sec) and Video with Audio (M = 739.6 sec) which were not significantly different from each other, but were both faster than Audio mode alone (M = 1490.3 sec). These results are displayed in Figure 4.13.

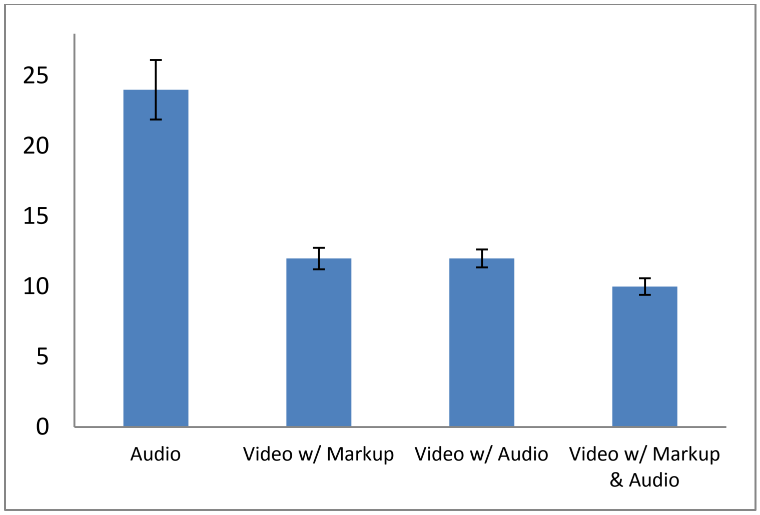


Figure 4.13: Mean completion times for each of the four experimental conditions. Error bars are standard errors

The total amount of data transmitted from Helper to Worker was compared across experimental conditions to evaluate the effectiveness of the collaborative tools in conveying adequate information to the Worker to complete the task successfully. A statistically significant main effect was found between four experimental conditions, F (3, 42) = 97.59, p< .01. Post hoc task found that the Helper used the least amount of transmitted data to complete the task in the Video with Markup (M = 1.99 MB) condition. This data usage amount was significantly less than that used in the Video with Markup and Audio mode (M = 5.90 MB), which was less than the amount used in Video with Audio mode(M = 12.75 MB), which in turn was less than Audio alone (M = 23.16 MB).

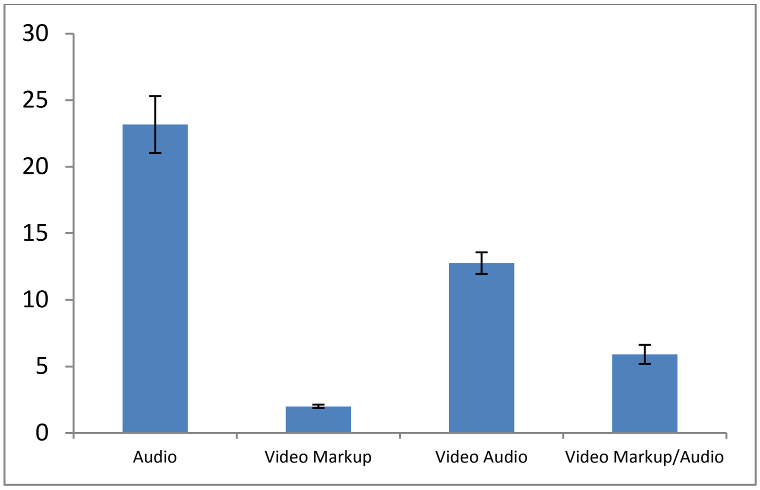


Figure 4.14: Mean data sent from Helper to Work (MB) for each condition. Error bars are standard errors

In regard to participants' perceived mental workload for completing the task with the different collaborative tools, it was found that ratings of global NASA-TLX scores were significantly different across various conditions, F (3, 42) = 12.2, $p < .01$. Post hoc test found that participants rated the Audio (M = 79.8) as the most mentally demanding condition, and Video with Audio (M = 61.7) and Video with Markup and Audio (M = 55.6) as the least demanding and not significantly different from each other.

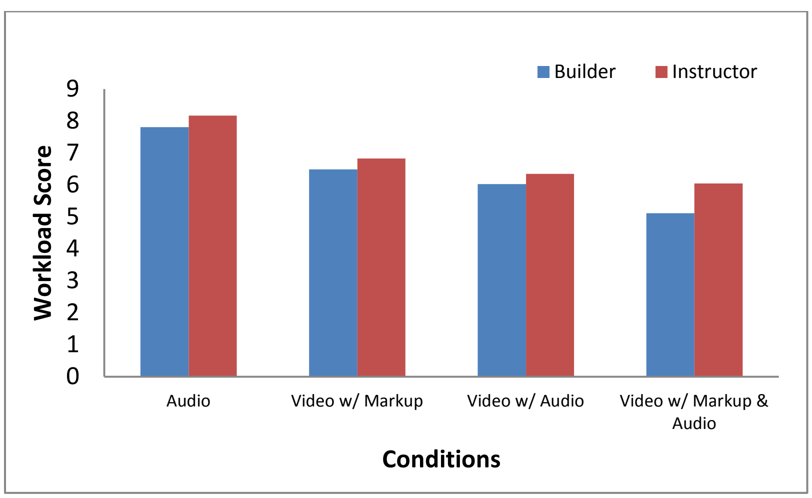


Figure 4.15: Mean TLX for each of the four experimental conditions.

4.4 Medical Demonstration

An evaluation relevant to the medical field was performed using the mobile application. The demonstration required participants to transfer points from an image containing desired point locations onto the surrounding tissue of open wounds. This task was selected to assess the precision and accuracy that the mobile application affords the Worker when applying reference data onto an object. The application of such an evaluation could prove that medical novices are capable of performing lifesaving emergency medicine under the guidance of a medical expert.

4.4.1 Participants

Six military participants volunteered for this demonstration (2 men and 4 women) ranging in age from 23-26 (M = 25) years. All participants had normal hearing and normal or corrected-to-normal vision.

4.4.2 Experiment Design

A within-subject design that was balanced using a Latin-square procedure was employed with three levels of interface (Paper, Side-by-Side, and Guide). All participants took part in a training session to familiarize themselves with the task and devices. The participants trained by performing point transferring practice trials per each experimental condition. Participants were given the option for more practice trials; however, none of them felt the need for more. The three experimental conditions were randomized per participant.

4.4.3 Apparatus

A mannequin with simulated soft tissue damage to its mid torso was used in the experiment. The mannequin's abdominal cavity was exposed showing a 10 inch x 10 inch section of synthetic skin with a variety of open wounds. The mannequin was positioned horizontally on its back on top of a flat table approximately waist high. Participants used a Samsung Galaxy Tablet which ran the developmental Android application to interact with the mannequin's wounds. The Galaxy Tablet was mounted on a stand to allow the participants to freely use their hands, as seen in Figure 4.16. The participants used a paint pen to mark the artificial skin with the point locations received per experimental condition.

Figure 4.16: Medical Demonstration Apparatus

4.4.4 Procedure

For each condition, participants transferred 33 dots from a reference image to a patch of synthetic skin on a mannequin.   They were asked to complete the task as fast as possible without making any errors. Three interface conditions were investigated: Paper, Side-by-Side, and Guide. In the Paper condition, participants used a printed image of the wound that showed the reference image dots to transcribe onto the mannequin, as seen in Figure 4.17 (a). The reference image was secured to cardstock and the physical dimensions of the printed image were the same as the digital image presented on the mobile device. The participants were not instructed nor restricted on how to hold the printed image. In the Side-by-Side condition, participants used the mobile device to retrieve wound reference image dots. The Side-by-Side interface displayed both a live perspective of the mannequin's wound section adjacent to a reference still image that showed the marks to transcribe, as depicted in Figure 4.17 (b). The participant could look through the mobile device by using the live perspective and/or could choose to look around the mobile device to apply the desired dots. In the Guide condition, participants likewise used the mobile device to retrieve wound dots. The Guide interface fused the live perspective of the wound section with the still image containing reference dots. A transparency value of 50 percent was applied to the still image so that the participant could interact through the image to apply the dots on the mannequin, as shown in Figure 4.17 (c).

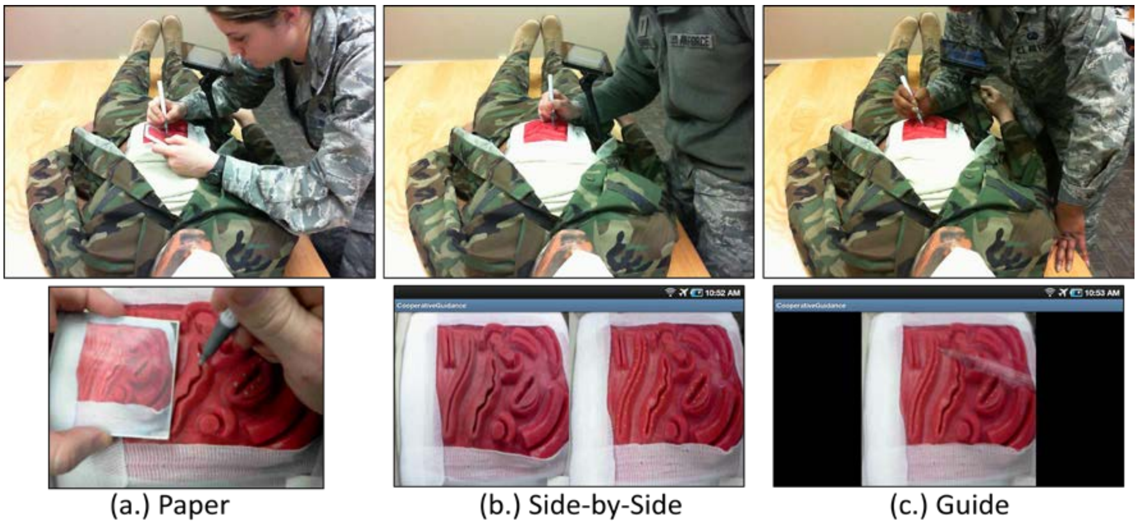


Figure 4.17: Interface conditions

4.4.5 Results

A three condition repeated measures Analysis of Variance (ANOVA) of these data revealed a statistically significant main effect for conditions, $F\ (2, 10) = 10.09$, $p< .05$. A subsequent post hoc Tukey-test with alpha set at .05 revealed that participants were significantly more accurate in their dot placement in the Guide condition then both Paper and Side-by-Side which were not significantly different from each other.

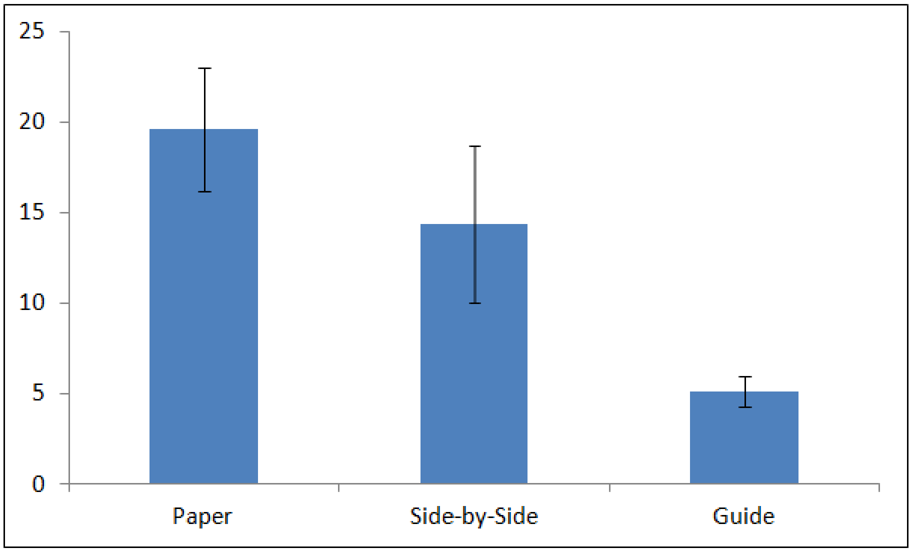


Figure 4.18: Mean error from Truth for each of the three experimental conditions. Error bars are standard errors

4.5 Power Assessment

A power assessment was performed on the developmental Android application to determine power costs the implemented communication capabilities have on the mobile device's battery. Knowing of the power effects for each of the capabilities, an informed determination of their necessity in supporting the remote collaboration session can be assessed and an informed decision on whether or not the remaining battery life can survive the duration of the task is capable. For example, if the battery was running low and the collaborating pair was utilizing 30 fps image sharing and if dropping the frame rate to 16 fps would not hinder cooperative performance, the pair could negotiate changing the fps to prolong the battery run-time, thus enabling extended communication. The power assessment conducted on the Android application yielded unique power profiles for each implemented communication capability.

4.5.1 Design

To begin determining the power cost of the communication capabilities, a Baseline power profile of the mobile device powered on and in a minimal idle state was collected. The Baseline power profile without any features activated was used to quantify the power costs of running the various communication capabilities on the battery. The power profiles captured for each implemented communication capability were compared to the Baseline configuration to distinguish their individual effects. The Baseline consisted of the mobile device turned on with a static display on the screen, not refreshing, and the screen time out disabled. Additionally, all wireless interfaces (e.g. Bluetooth, WiFi, etc.) were disabled, and no integrated devices (e.g. camera, speaker, etc.) were used. Moreover, the device was configured so as not to go into sleep mode. The Baseline condition was representative of the minimum idle state that the mobile device can be in while powered on.

4.5.2 Apparatus

To account for variation in performance between different mobile devices of the same model, three Samsung Galaxy Tablets were utilized in the recording of power effects for each of the communication modalities. Running on the mobile devices was the designed power measuring process (PMP) that was used to record and log various run-time settings of the mobile device and its battery. In addition to the mobile devices, three Gateway laptops were used in the assessment of network communication power effects, serving as remote hosts echoing network traffic from the mobile devices. Also, a Linksys 2.4 GHz wireless-G broadband router was used in enabling the wireless local area network.

4.5.3 Procedure

The generation of power consumption profiles associated with the implemented communication capabilities supporting remote collaboration required identification of the mobile device's hardware components used for each capability. The mobile device hardware components used were the display, speakers, microphone, network interface card, and camera. In addition to the hardware components, three pre-determined fidelity usage levels (High, Medium, & Low) were examined for the components that had dynamic ranges.

Isolated hardware components and respective fidelity levels were executed on fully charged batteries and ran until the battery was fully depleted and the mobile device turned off. Several power measurements (3-4) for each identified hardware component and fidelity level were performed to capture power trends. Moreover, the measurements were run in a climate controlled temperature of 68-72 degrees Fahrenheit.

The Baseline condition's three fidelity levels corresponded to the brightness level of the non-refreshing screen (High – 255, Medium – 127, Low – 0). For assessing the power used by a refreshing display, the mobile device was configured similarly to the Baseline condition with the exception of the display's ability to refresh. The display's refresh toggled between solid white and blue screens as fast as the mobile device would permit with the varied screen brightness fidelity levels (High – 255, Medium – 127, Low – 0). Audio power usage was determined by setting the mobile device into Baseline fidelity level 0 condition and playing a continuous wave file at various volume levels. The fidelity levels associated with audio were High – 16, Medium – 8, and Low – 1. Network power usage was captured in two ways. The first was the power associated with the WiFi hardware powered on and connected to a network without transmitting or receiving network traffic. The second was connected and transmitting and receiving network traffic at three fidelity levels (High – 622KB/s, Medium – 342KB/s, Low – 172KB/s). The camera and microphone power usages were assessed while powering the hardware components.

4.5.4 Results

Power consumption of the communication capabilities were isolated and analyzed to compile power profiles. These profiles can be used to determine dynamic runtime conditions in orderto prolong the mobile device's battery duration. Additionally, the communication capabilities status messagecan activate and deactivate unnecessary or unused capabilities.

The power measurements captured for the Baseline condition were analyzed against time (ms) and battery state of charge, as seen in Figure 4.19. The data for the High fidelity level (255) revealed a linear equation of $y = -4e^{-06x} + 100.42$ and an R-square value of 0.9995. The Medium fidelity level (127) resulted in a linear equation of $y = -3e^{-06x} + 99.712$ and an R-square value of 0.9998. The Low fidelity level (0) showed a linear equation of $y = -2^{e-06x} +99.749$ and an R-square value of 0.9999.

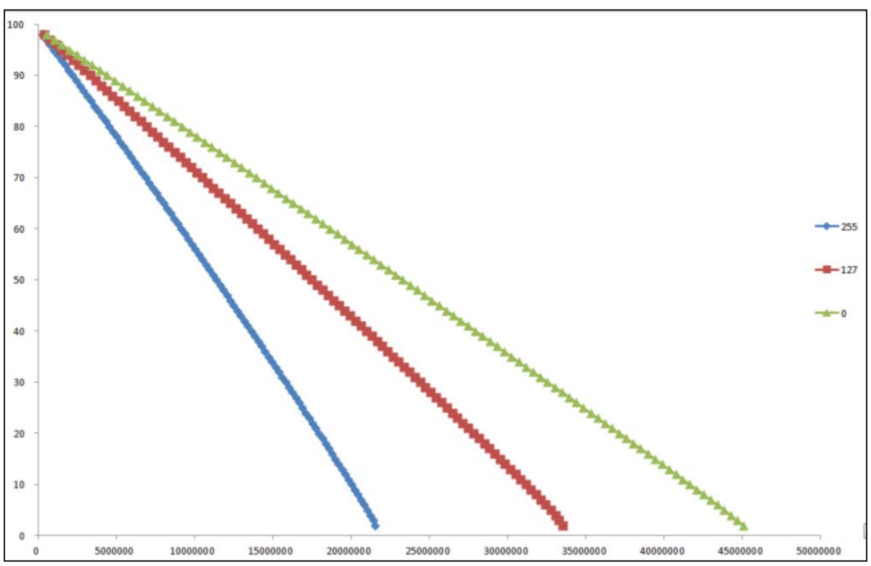


Figure 4.19: Average Baseline power profile per fidelity levels

The power measurements captured for the Display condition were analyzed against time (ms) and battery state of charge, as seen in Figure 4.20. The data for the High fidelity level (255) revealed a linear equation of $y = -5e^{-06x} + 100.55$ and an R-square value of 0.9994. The Medium fidelity level (127) resulted in a linear equation of y = $-3e^{-06x}$ + 100.02 and an R-square value of 0.9996. The Low fidelity level (0) showed a linear equation of y = $-3e^{-06x}$ +100.03 and an R-square value of 0.9998.

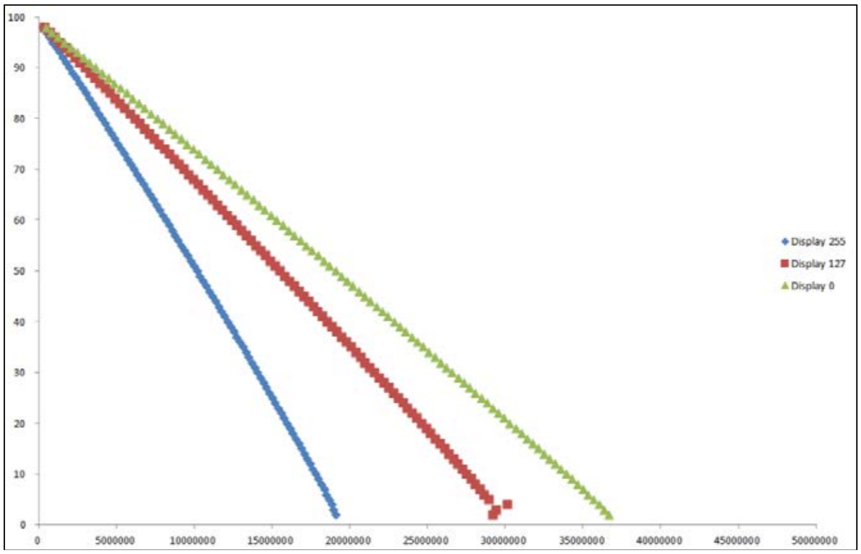


Figure 4.20: Average display power profile per fidelity levels

The power measurements captured for the Audio condition were analyzed against time (ms) and battery state of charge, as seen in Figure 4.20. The data for the High fidelity level (15) revealed a linear equation of y = $-3e^{-06x}$ + 99.915 and an R-square value of 0.9999. The Medium fidelity level (8) resulted in a linear equation of y = $-2e^{-06x}$ + 99.983 and an R-square value of 0.9999. The Low fidelity level (1) showed a linear equation of y = $-2e^{-06x}$ +99.783 and an R-square value of 0.9999.

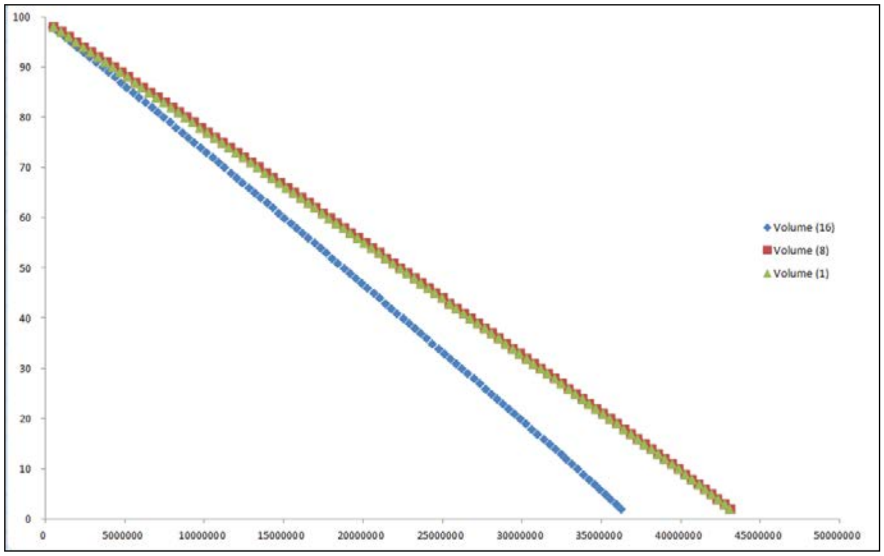


Figure 4.21: Average audio power profile per fidelity levels

The power measurements captured for the Network condition were analyzed against time (ms) and battery state of charge, as seen in Figure 4.20. The data for the High fidelity level (622KB) revealed a linear equation of $y = -3e^{-06x} + 98.95$ and an R-square value of 0.9999. The Medium fidelity level (342KB) resulted in a linear equation of $y = -3e^{-06x} + 99.312$ and an R-square value of 0.9998. The Low fidelity level (172KB) showed a linear equation of $y = -3e^{-06x} +99.571$ and an R-square value of 0.9999.

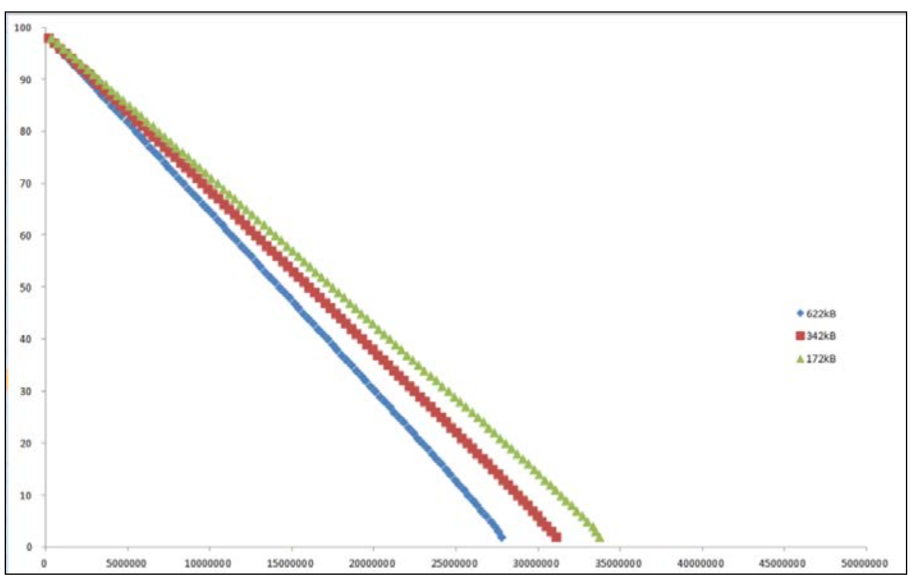


Figure 4.22: Average network power profile per fidelity levels

5. DISCUSSIONS

There were three objectives for this body of research:

1) Validate mobile device usefulness for distributed collaboration
2) Explore the relative effectiveness of both the human users and the mobile devices when multimodal communication capabilities are presented under remote supervisory guidance.
3) Produce an effective power consumption adaptive Android algorithm that can dynamically adjust the device contexts and presentations to ensure that the battery charge survives the entire task.

The following sections will elaborate on these objectives and explain how each of them was accomplished in this research initiative.

5.1 Mobile Device Usefulness

To evaluate the usefulness of mobile devices utilized for remote collaboration on physical tasks, three key areas were assessed: ability to perform at least as well or better than traditional CSCW systems, mobility, and task completion. Traditional CSCW systems facilitate remote collaboration by enabling distributed partners to communicate through a variety of communication capabilities. The most commonly used communication mediums in CSCW systems are visual and auditory capabilities. Visual information sharing in the form of streaming video, still images, and annotation markups are the primary usages. Streaming video and acquisition of still images are achieved through the use of video capture devices that are controlled by either member of the collaborating pair. For example, Kraut (2003) utilized a head mounted camera to share perspective awareness between Workers and Helpers using an affixed camera on the Worker's head. Kuzuoka (1992) used a static mount with a motorized gimbaled camera that was able to sweep across the workspace of the Worker enabling the Helper to monitor the task progress. Kirk (2002) used an overhead stationary camera to capture the Workers' and Helpers' actions and physical tasks.

Regardless of video capture capability, the need for visual information sharing is of critical importance in remote collaboration between distributed individuals working together on a physical task. Mobile devices are well equipped to support this communication medium. It is the norm that mobile devices have integrated cameras. Additionally, mobile devices have the capability to connect to external video capture devices through a variety of wireless channels (Bluetooth, WiFi, ZigBee, etc.) in addition to physical input ports, such as USB. Like traditional CSCW systems, mobile devices use the TCP/IP network configuration to disseminate visual information between collaborating pairs.

Some CSCW systems add to the captured visual information by including graphical context in the form of annotations. Annotations most commonly used are predefined and free-form marks that are merged to the shared visual information. Annotations can serve to draw attention to a region within the captured visual image or illustrate procedural instructions to apply to the physical task. Ou (2003) used a touch screen interface and a stylus to generate the free-form annotations used for collaboration. Similarly, Fussell (2004) used real-time drawings added to streaming video to share visual information between distributed cooperative pairs working on a physical task. This white-boarding communication capability can be easily incorporated and controlled by mobile devices. Touch screens are quickly becoming the standard input interface mechanism for mobile devices, and the use of a finger and/or stylus is common practice in notation. Real-time editing of captured still images with graphical annotations, as well as fusing markups and streaming video, is achievable through readily available graphical libraries for mobile devices. It can be argued that the mobile device's form factor improves this capability over traditional CSCW systems.

Mobile devices are not restricted in movement or confined to table-top setups as most traditional CSCW systems are due to power and network interconnectivity cabling. Additionally, a mobile device's orientations can support dynamic movement in three-axes, whereas CSCW systems are mostly static in placement. This ease of use enhancement promotes a more natural markup editing, similar to artists and their composition pads.

Auditory information sharing for CSCW systems is often done through the use of audio capture devices and a network connectivity to transmit audio signals between distributed pairs working on a physical task. CSCW systems incorporate microphones, speakers, headsets and other audio input/output technology to support audio communication. Unlike traditional CSCW systems, mobile devices' lineage started with audio communication, since the first mobile devices stemmed from mobile telephone services. CSCW uses TCP/IP to transmit auditory information, whereas mobile devices can be configured to transmit audio signals through a combination of cellular, TCP/IP and other RF means.

A clearly distinguished advantage that mobile devices have over traditional CSCW systems is the ability to be carried on the person and into various environments and situations where traditional CSCW systems cannot perform. Therefore, mobility was assessed as a vital contributing factor in evaluating the usefulness of mobile devices in remote collaboration. Mobility affords collaboration with distributed parties anywhere and at any time. Mobile devices support mobility through the use of various built-in communication channels that are transparent to the user. It is often the case in today's rapidly moving distributed workforce that an individual faces a task that is outside of their expertise while on a remote job site, traveling between locations, or attempting to respond to an unplanned event. Mobile devices have the ability to reach out and communicate with experts whose assistance could prove critical to the overall completion of the task at hand.

Mobile devices enable on-the-move processing of information and can leverage a variety of built-in capabilities to capture the mobile device's surroundings. Additionally, mobile devices enable on-demand retrieval and communication which now seamlessly integrates into peoples' lives. Mobile devices are so proliferated in today's society that individuals needing remote assistance already possess the power of CSCW capabilities usually at hand.

The third factor assessed was task completion, which is equally important as device mobility and similarity to traditional CSCW systems. If the cooperative tasks could not be accomplished through the use of mobile devices, then obviously mobile devices would not to be an ideal tool for remote collaboration. This factor was assessed under a variety of relevant scenarios to explore the versatility that mobile devices have in distributed task completion. The scenarios evaluated were IED disposal, a find, fix, and tag task, a building block assembly, and a medical care task. Of the 52 participants utilizing a mobile device on tasks presented to them while communicating with a remote assistant, all 52 were successful in completing their objectives. Additionally, the participants required minimum training to utilize the mobile devices on the tasks as they all had previous exposure to mobile devices outside of experimental conditions. Aiding in the task completion was the intuitive information portrayal that the mobile device facilitated between Workers and Helpers.

The evaluation of the assessed factors supports the conclusion that mobile devices can effectively enable and contribute to remote cooperative pairs working on a physical task. Therefore, mobile devices are in fact extremely useful in distributed collaboration.

5.2 Human and Mobile Device Effects

The second component of this dissertation was the exploration of the effects that multimodal communication, as presented on a mobile device, has on the human participants as well as the mobile devices while participating in remote collaborations. Section 5.2.1 will analyze the impact multimodal communication capability has on the human user, and section 5.2.2 will explore the effects on the mobile device in terms of power usage and use adaptability in various situations.

5.2.1 Human Performance Effect

In regards to human performance, Wickens and McClarley's research (2008) found that systems and interfaces utilizing multiple modalities are more advantageous to the user then those that do not have those capabilities. Multimodal research findings suggest that multimodal interfaces allow users to process different modality information concurrently with better cognitive understanding of the task. Moreover, the presentation of multimodal information serves well in cognitively demanding environments that require Workers to share their cognitive focus and attention across several complex and concurrent events. Using the empirical data collected from the various cooperative scenarios, we can assess the impact that multimodal communication executed on a mobile device has on the human in terms of workload, performance time, conversational strategy, accuracy, and confidence.

5.2.1.1 Workload

The NASA-Task Load Index (TLX) is one of the most effective and widely used measures of perceived mental workload currently available (Farmer & Brownson, 2003; Nygren, 1991; Wickens & Hollands, 2000). It assesses six sources of workload: Mental Demand, Temporal Demand, Physical Demand, Performance, Effort, and Frustration to provide a global workload rating on a scale of 0 to 100 (Nygren, 1991). The six workload sources are then combined to form an overall workload index on a scale of 0 to 9. Figure 5.2 shows the workload results for the most complex task studied, the block assembly, which required substantial Worker and Helper interactive communication for completion. The results show that as the collaborating pairs utilized more modalities their respective workloads decreased.

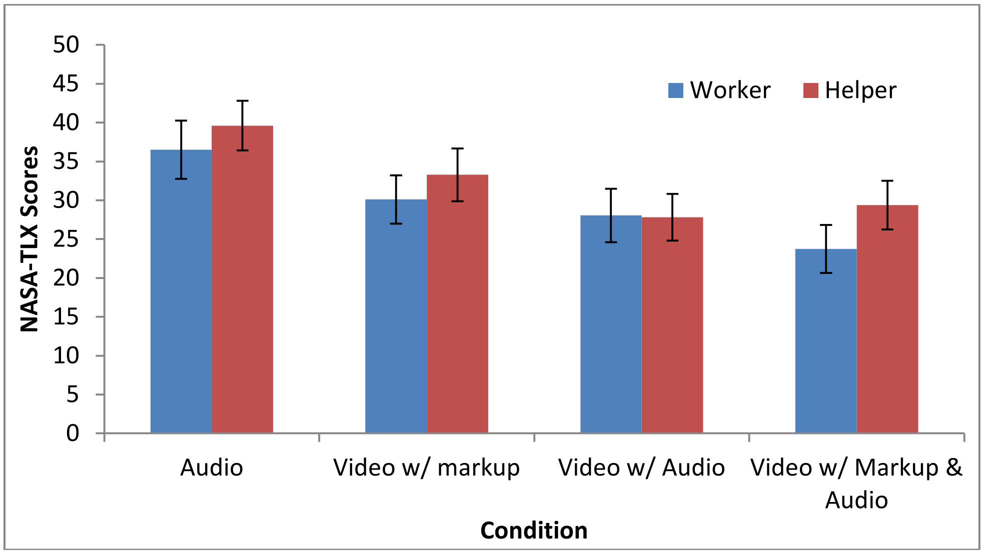


Figure 5.1: Workload TLX for Block Assembly Task

5.2.1.2 Performance Time

The findings from the cooperative task evaluations showed that performance times decrease when Workers utilize more modalities while communicating with remote Helpers. Workers were faster at completing each of the three evaluated tasks, as shown in Figure 5.3, when using the most modality condition, Video with Markup and Audio, than any other multimodal combination (Audio, Video with Mark, and Video with Audio). In addition to decreasing performance times of the Worker, the Helper's performance times were reduced when using more modalities to relay instructional information. The Helper spent less time explicitly describing task objects and was more efficient in providing supervisory guidance when leveraging multimodal communication capabilities as opposed to when single communication modalities were used for task completion.

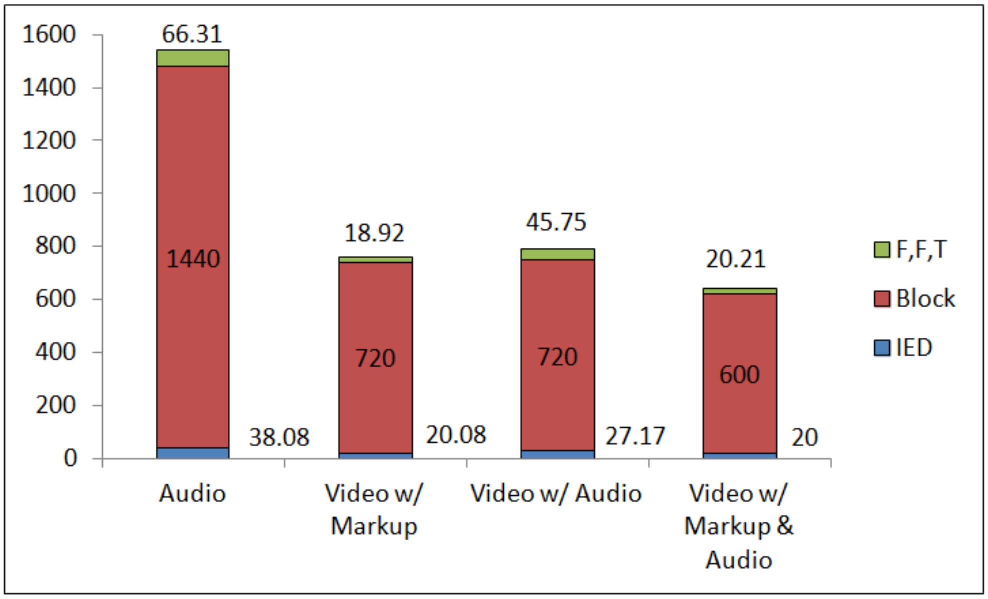


Figure 5.3: Completion times (sec) for remotely assisted tasks

5.2.1.3 Conversational Strategy

Communication strategy between Workers and Helpers was positively affected by the use of multimodal communication capabilities. When comparing the Video with Audio and Audio only conditions, the use of shared visual perspective resulted in a faster convergence of understanding, as well as having an impact on the style of the verbal directives. For example, in the Audio only condition, with no shared visual information, the remote Helper's verbal directives were much more descriptive in definingthe appearance of the task object. For example, in the Find, Fix & Tag task Helper's directives were as follows: "The first guy has no hat [pause] white beard [pause] and a gray shirt. The next guy has a brown hat [pause] small black beard [pause] and a white shirt." Alternatively, in the Video with Audio condition, the Helper's verbal directives provided contextual information on the task object's location in the shared visual field. In one such task using the Video with Audio mode, the Helper's comments were as follows: "all the way to the back next to the car [pause] that one [pause] yep", "the fifth one to the right"). Moreover, in this dual audio and visual mode, the remote Helper was able to use pronouns such as "that one", "him", "next one" to convey and direct the Worker's aim towards the correct target. The descriptive and contextual information experienced is similar to the classification of utterance ideas of *Referents* and *Position* presented in Kraut (2003).

Figure 5.3: Find, Fix, & Tag task involving shared visual information between Worker and Helper

5.2.1.4 Accuracy

The findings from the three tasks evaluated shows that when Workers use multiple modalities concurrently their accuracy performance improves, as shown in Figure 5.5. This finding is very apparent when collaboration is performed in visually complex or difficult to describe environments such as in the case of the Find, Fix and Tag task experiments conducted in this body of research.

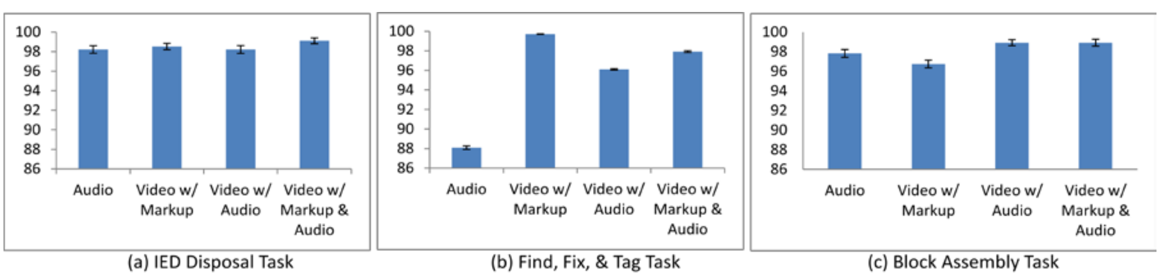


Figure 5.4: Mean accuracy results for Worker and Helper collaborating tasks

5.2.1.5 Confidence

Confidence is a key factor in any remote collaboration between Workers and Helpers. Timing and accuracy can be affected if either cooperative member questions the specifics of a guidance procedure or fails to perform the appropriate actions on the physical task. Exploring the effects that multimodal communications has on Worker's confidence, a subjective measurement of confidence in the information received as well as the Worker's resulting performance on a task was conducted using a seven point scale Liker questionnaire. Confidence questions asked were:

- How confident are you that you [*action*] the correct [*object*] every time?
- How confident are you in the information you received?
- How many [*object*] do you think you [*action*] correctly?

The *action* and *object* of the questionnaire were replaced with task specific roles and items per evaluated scenario. For example, in the IED use case *action* was replaced with "cut" and *object* was replaced with "wire". The results found that Workers' confidence improved in conditions where more communication modalities were used. Workers provided the highest confidence marks for the Video with Markup and Audio condition than for any other modality. The condition that that scored the lowest was the Audio only mode. The results for IED disposal and the Find, Fix, & Tag tasks are shown in Figure 5.6.

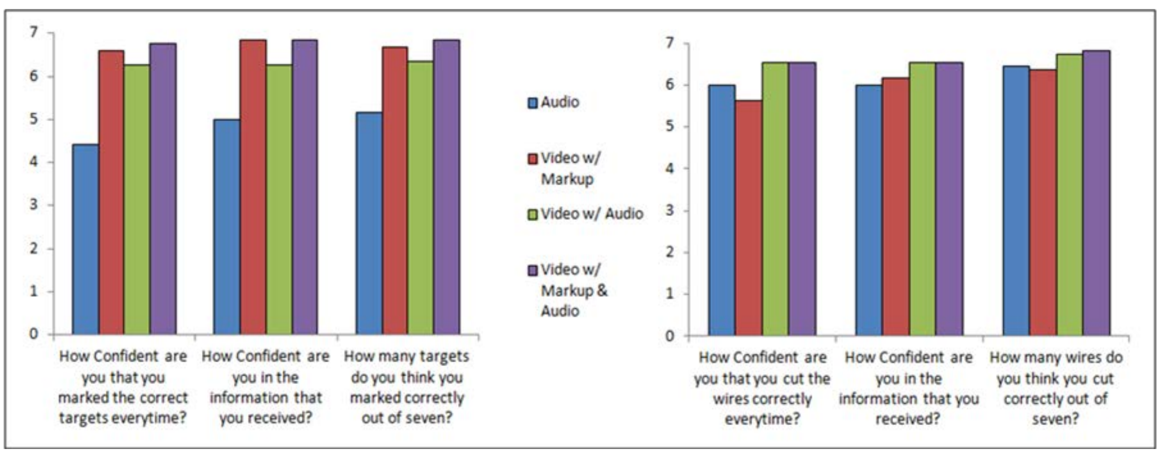


Figure 5.5: Workers' confidence scores for IED and Find, Fix, & Tag Tasks

5.2.2 Mobile Device Performance Effect

A primary consideration when using multimodal communications on mobile devices is the amount of information processing required in relaying task procedural guidance. Figure 5.5 shows the amount of data that the Helper was required to transmit to the Worker for (a) the Block Assembly and (b) the IED Disposal tasks. When a task required auditory information sharing, combining the auditory modality with the visual modality afforded a reduction in total information, measured in bytes, needed to successfully accomplish the task at hand. In the case of the Block Assembly task, the Helper was able to reduce transmitted auditory information by 58%; and in the case the IED Disposal task, audio was reduced by 15%. The savings are greater when the tasks are visually complex and difficult to describe requiring additional data sharing in regard to the current state of the task. Additionally, the data also reflects a reduction in the amount of required visual modality information when visual is combined with the auditory modality. The reduction of total modality information attributed to using multiple modalities concurrently has a positive impact on power consumption as less data needs to be transmitted, processed and presented through the mobile device's hardware components.

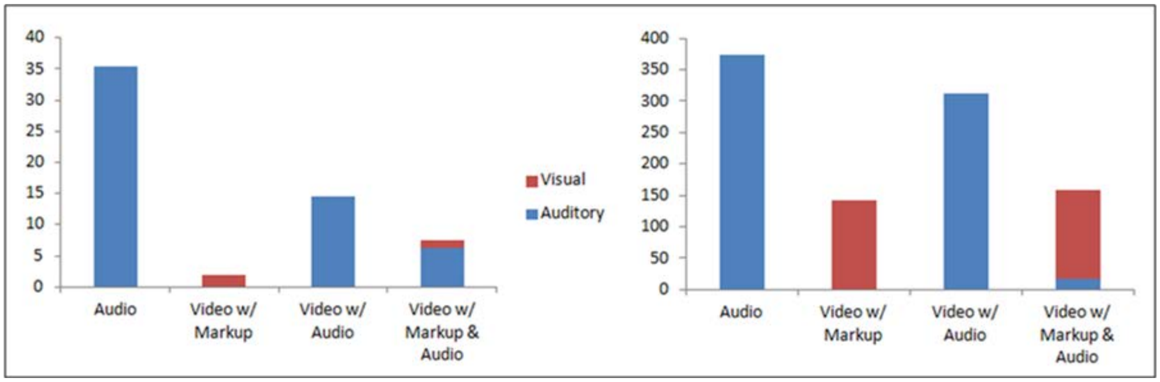


Figure 5.6: Helper data sent to Worker for IED Disposal and Block Assembly Task

5.3 Power

The third objective of this dissertation was to produce an effective power consumption adaptive algorithm that can dynamically adjust the mobile device's context and presentation to improve the likelihood that the battery charge survives the entire task. The ability to remotely monitor run-time power conditions and power consumption penalties per communication modality is a unique feature that the Android application developed for this research initiative possesses. This capability alleviates the burden of the Worker to monitor their device's battery state of charge solely as is the case for traditional CSCW systems. The sharing of the mobile device's current multimodal communication fidelity and the battery's current state of charge can be used to locally and remotely assess the power consumption of the active communication link between a Worker and Helper. The power profiles captured are used as inputs into the equation (2) to determine the active rate of battery charge consumption, which used in equation (1), can solve the T_r (time remaining) when B_p (current battery percentage of charge) is known.

$$B_p + (R_b)T_r = 0 \qquad (1)$$

$$R_b = r + a(\text{volume}) + n(\text{throughput}) + w(\text{radioOn}) + c(\text{cameraOn}) - 4b(\text{brightness}) \qquad (2)$$

B_p = Battery current percentage of charge
R_b = Rate of battery charge consumption per unit of time
T_r = Time remaining
r = Screen refresh consumption
a = Audio consumption as a function of volume (15, 8, 1)
n = Network consumption as a function of throughput (622kB, 342kB, 172kB)
w = Wireless NIC radio consumption (on/off)
c = Camera consumption (on/off)
b = Baseline consumption as a function of brightness (255, 127, 0)

Knowing T_r (time remaining) the Worker and Helper can negotiate communication capabilities if T_r is not sufficient enough to complete the collaborative task. Additionally, the ability to dynamically change modality and fidelity settings locally or remotely can prolong the operational duration of the mobile device when used as a tool for cooperative interaction.

6. CONCLUSION

This chapter revisits and summarizes the main research objectives of this dissertation. Additionally, it highlights original contributions and potential future research in this area. In the first section, *Research Summary,* the objectives are articulated and a summary of how each objective was successfully accomplished is explored. The second section, *Original Contribution*, reports the original contribution made in this research as well as benefits and practical applications learned from the investigation. Lastly, in the section *Future Work*, the chapter discusses potential future work which could further this research endeavor and expand on the already made contributions to computer supported cooperative work (CSCW) and mobile power management research.

6.1 Research Summary

This dissertation investigates and demonstrates the effects that various modalities have on remote collaboration between individuals utilizing interactive mobile devices with respect to human performance and power consumption. The research documents the design and implementation of an Android interactive communication suite that supports multimodal communication capabilities facilitating remote collaboration. The original developed software features were designed to enhance human performance through mobile on-demand, ease of use, intuitive multi-touch interfaces and configuration menus. The mobile software additionally supports dynamic information sharing through various mobile networks and peripheral connectivity. Moreover, the software permits real-time changes so that the mobile device user can leverage of the most appropriate presentation mode seeking to maximize their effectiveness in the current mobile surroundings. A series of experiments and demonstrations using the system were executed to explore the human performance effects and mobile device utilization during performance of cooperative, distributed objectives. The experiments included scenarios that are particularly valuable to the military first responders in field operations. Additionally, an investigation was conducted focusing on mobile device power consumption and conservation. Remote Helpers and Workers could negotiate and simultaneously monitor power consumption effects of active multimodal communication capabilities striving to ensure that power consumption and battery life does not prematurely expire prior to task completion.

Three research objectives were addressed and successfully accomplished in this dissertation adding value to CSCW research and mobile power management. The first objective was to validate mobile device usefulness for distributed collaboration. The second objective was to explore the relative effectiveness of both the human users and the mobile devices when multimodal communication capabilities were presented under remote supervisory guidance. The third objective was to produce an effective power consumption adaptive Android algorithm that can dynamically adjust the device contexts and presentations to ensure that the battery charge survives the entire task.

Assessing the usefulness of a mobile device in remote collaboration scenarios was achieved through various applied experiments and demonstrations. Factors analyzed determining usefulness were mobility, task completion, and performance compared to traditional CSCW systems. Results from the experimentation and analysis support the finding that mobile devices are in fact useful for remote collaboration. The effects of multimodal communication on human performance and on a mobile device were likewise evaluated in applied experiments and demonstrations. In regard to human performance, it was proven that the use of multimodal communication capabilities resulted in improved participant performance when compared to single modal communication in the analyzed scenarios executed in this dissertation research. The empirical performance data collected, including task completion times, accuracy, user workload and confidence measurements, produced findings that are comparable to other research initiatives assessing human performance while using multimodal communications; however, not in the mobile context as evaluated in this body of research.

In regard to mobile device performance, the use of multiple modality capabilities when used together showed a reduction in the amount of data used by a single modality when used in isolation. The development of an effective power consumption adaptive algorithm was accomplished in the following stages. First, the isolated communication power consumption rates of each multimodal communication capability tested was quantified to gain an understanding of how each device modality is affected during runtime. The next component in determining the power conservation effect was achieved by measuring various combinations of shared communication capabilities in regard to their comparative power usage. It was found that with the knowledge of power consumption effects and knowing which communication capabilities were active, the remote partners cooperating on a task could adjust mobile device settings to prolong battery life.

Chapters 3-5 highlight the details of the above mentioned research. Chapter 3 details the design and implementation of various multimodal mobile communication capabilities, as well as power measurement and status control messaging. Chapter 4 reports on the experiments and demonstrations conducted and their respective findings. Chapter 5 discusses in detail the effects multimodal communication has on task completion times, accuracy, workload, and user confidence.

The findings of this dissertation research have significant implications for the design, deployment and development of future mobile collaborating infrastructure applications for both military and civilian use. Results from the experiments and demonstrations show that mobile devices can increasingly support the communication capabilities necessary to successfully complete tasks jointly performed by a Worker and a remote Helper. Additionally, the ability to share power consumption rates between collaborating individuals enables more efficient power usage through the toggling of information sharing capabilities and alleviates the burden being solely on the remote mobile device user. This investigation has provided justification for further development of mobile multimodal collaborative applications for distributed military or civilian first responders. The documentation of these study findings has successfully met the research objectives outlined in this dissertation.

6.2 Original Contribution of Research

This dissertation adds to the body of work exploring and understanding the impact that mobile devices have on CSCW as well as mobile power management. Previous CSCW research has demonstrated the usefulness of multimodal communication capabilities executed on PCs in a static setting; however prior research has not investigated the performance impact in a mobile domain. The original contribution performed in this body of research was an assessment of team collaboration leveraging newly developed multimodal communication capabilities in a mobile capacity. Contributing software advancements include a unique power measurement process and energy profiling capability, real-time exchange protocol for the modification of streaming information and device settings, along with enhanced audio/visual mobile presentation software. The empirical data gathered and resulting analysis provides a further understanding of the relative effectiveness of various mobile device communication modalities when used by a team of individuals engaged in remote collaboration. Additionally, the exploration and experimentation developed in this research addresses both static and dynamic interactions, as highlighted in the IED disposal and Find, Fix, and Tag experiments in Chapter 4. The use of a mobile device to remotely assist in task completions involving the manipulation of physical objects and interaction with distant surroundings provides further contribution to CSCW research.

Remote power monitoring is an original contribution of this dissertation. Battery power research has failed to achieve and explore how distributing local power consumption rates to a remote party may improve team communications and interactive cooperation. This research has identified that power savings can be achieved when teams jointly monitor the finite power supply of a mobile device simultaneously and negotiate dynamically the modification of communication capabilities to conserve power. This research designed and developed a software measurement process utilizing custom power equations that calculate expected run-time remaining based on a mobile device' active communication dissemination and local presentation modes. The derived power performance values of the battery were packaged and shared to all cooperative collaborating members for greater remote power awareness. The power equations were incorporated into the mobile software suite permitting collaborating pairs the ability to assess and monitor simultaneously the real-time power consumption of the mobile device. With the knowledge of how a mobile device processes communication in regard to power consumption rates, team members can determine remaining runtime using current communication modes and can adjust accordingly to prevent premature battery depletion prior to the completion of the joint objective.

An additional contribution implemented in this dissertation is a remote software protocol used for the modification of the mobile device's presentation mode and output settings. The value that this feature supports is real-time adjustment to disseminated data from the mobile device. In addition to conserving power, this feature also permits remote collaborating parties the ability to increase or reduce resolution and fidelity to maximize the ease of communication. Remotely adjusting local device settings, such as transmitted frames per second of captured video, facilitates remote parties in obtaining a higher degree of situation awareness as well as positively affecting the communication between the interacting pairs. The software protocol was designed to scale for future expansion to incorporate new mobile device features, yet to be added, that would benefit from team remote control. The protocol leverages tag fields similar to extensible markup language (XML) messages that are assigned to particular mobile device features. Tags can have associating resolution and fidelity values appropriate with mobile device features that are adjustable. For example, screen brightness has a tag field with a luminous value that can be set within the range of $0 \leq X \leq 255$.

A technology-based contribution made in this dissertation is real-time dynamic adjustable transparency of shared still images. This feature allows mobile device users to apply remote guidance "on-top-of" the live perspective for improved application of directives. The value of this contribution allows mobile device users to interact with the local scene while "looking through" the instructional directives received from a remote individual or system. (Similar transparency concepts are just recently being introduced into commonly used consumer electronics products.) Leveraging touch inputs from the user, visual information's transparency can be adjusted from full transparency (invisible) to any user driven partial transparency value. The use of gesture motions is conducive for mobile device interaction supporting quick changes to the transparency effect in a non-interference method that does require on screen restate.

6.3 Future Work

There are numerous future work initiatives that can be done to further this body of research. Examples include the following.

- Investigate the effects that dual cameras, embedded into mobile devices, could have in enhancing perspective sharing. For example, 3D images could be used to capture an object requiring manipulation. Additionally, visual perspective sharing could be expanded to capture an immersive scene through 3D environment stitching (from a series of tiled still images) supplying remote helpers with a total virtual awareness of the surrounding environment, as well as the object of interest, as shown in Figure 6.1.

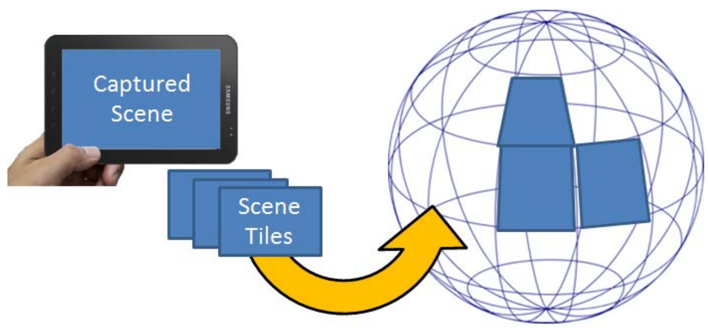


Figure 6.1: Immersive 3D scene generated from a series of still images

- Extend the markup annotation capability to include persistent memory of regional markup information in the context of an immersive scene. Simply put, remote annotations could be archived and displayed only when the mobile device's field of view overlaps with the region containing markup of information (similar to augmented reality and icon placement).
- Integrate these (and similar) capabilities into "heads-up see-through" display, such as Goggle Glasses.
- Further evaluation of communication between multiple teams of individuals utilizing different configurations of workstation and mobile devices is warranted. Exploration into an enhanced communication infrastructure that supports multiple users simultaneously on both sides (Workers/Helpers) to collaborate and inject expertise in the shared space is a research area that could be expanded. Additionally, the ability to

toggle between the various perspectives the multiple individuals share could impact power, bandwidth, and human performance.

- Assess 3D audio executed on a mobile device supporting spatialized separations of auditory information from multiple sources to evaluate performance benefits.

Mobile devices are quickly becoming a permanent fixture in individuals' daily activities. Maximizing the potential benefits of these mobile computing devices decidedly improves user experience and productivity. In particular, regarding the focus of this dissertation's research, the enhanced capabilities of mobile device usage greatly facilitate the cooperative efforts of physically separated individuals in the completion of any number of specialized tasks.

REFERENCES

1. Ballinger, B. Allauzen, C., Gruenstein, A., & Schalkwyk, J. (2010). On-demand language model Interpolation for mobile speech input. In: *Proceedings of ISCA 2010*, Makuhari, Chiba, Japan, 1812-1815.

2. Brennan, S.E. (2004). How conversation is shaped by visual and spoken evidence. In: J. Trueswell & M. Tanenhaus (Eds.), *World Situated Language Use: Psycholinguistic, Linguistic and Computational Perspectives on Bridging the Product and Action Traditions.* Cambridge, MA. MIT Press.

3. Brewster, S., & Brown, L. (2004). Non-visual information display using tactons. In: *Ext. Abstracts CHI 2005*. ACM Press, 787-788.

4. Buck, R., & VanLear, C.A. (2002). Verbal and nonverbal communication: Distinguishing symbolic, spontaneous, and pseudo-spontaneous nonverbal behavior. *Journal of Communication,* 52(3), 522-541.

5. Carrol, A. (2010). An analysis of power consumption in a smartphone. In: *Proceedings of 2010 USENIX Conference.* CA, USA.

6. Chang, A.,& O'Sullivan, C. (2005). Audio-haptic feedback in mobile phones. In: *Ext. Abstracts CHI 2005.*AMC Press, 1264-1267.

7. Clark, H.H., & Brennan, S.E. (1991). Grounding in communication, In L. B. Resnick, R. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition,* 127-149, Washington, DC: APA.

8. Clark, H.H., & Krych, M.A. (2004). Speaking while monitoring addresses for understanding. *Journal of Memory & Language*, 50(1), 62-81.

9. Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

10. Dong, M., & Zhong, L. (2011). Self-constructive high-rate system energy modeling for battery-powered mobile systems. In: *Proceedings of International Conference of Mobile Systems, Applications, and Services,*335-348

11. Farmer, E., & Brownson, A. (2003). Review of workload measurement, analysis and interpretation methods. In: European Organization for the Safety of Air Navigation (Eurocontrol): Report for Eurocontrol Integra Program. Brussels.

12. Flinn, J., &Satyanarayanan, M. (1999). Energy-aware adaption for mobile applications. In: *Proceedings of the 17th ACM symposium on Operating systems principles.* 33(5).

13. Fussell, S.R., Kraut, R.E., & Siegel, J. (2000). Coordination of communication: effects of shared visual context on collaborative work. In: *Proceedings of CSCW 2000*, NY: ACM Press, 21-30.

14. Gale, S. (1990). Human aspects of interactive multimedia communication. In: *Interacting with Computers*, 2(2), 175-189.

15. Gaver, W. (1997). Auditory interfaces, In: Helander, M., Landauer, T. & Prabhu, P. (Eds), *Handbook of Human-Computer Interaction*, Amsterdam: Elsevier Science.

16. Gergle, D., Kraut, R.E., &Fussell, S.R. (2004). Action as language in a shared visual space. In: *Proceedings of CSCW '04.* 6(3), NY: ACM Press. 487-496.

17. Gergle, D. (2005). The value of shared visual space for collaborative physical task. In: *Proceedings of CHI '05*, 1116-1117.

18. Heo, S., & Lee, G. (2011). ForceTap: extending the input vocabulary of mobile touch screens by adding tap gestures. In: *Proceeding of MobileHCI 2011*, Stockholm, Sweden. 113-122

19. Hoggan, E., Crossan, A., Brewster, S., & Kaaresoja, T. (2009). Audio or tactile feedback: which modality when? In: *Proceedings of CHI 2009.*AMC Press, Boston, MA.

20. Isaac, E.A, & Tang, J.C. (1993). What video can and can't do for collaboration: a case study. In: *Proceedings of AMC Multimedia '93.* New York: ACM Press. 199-206.

21. Kirk, D., Crabtree, A., & Rodden, T. (2005). Ways of the hand. In: *Proceedings of ECSCW 2005*, 18(22), Paris, France, 1-22.

22. Kirk, D., Rodden, T., &Fraser, D.S., (2007). Turn it this way: ground collaborative action with remote gestures. In *Computer Human Interaction (CHI): Distributed Interaction*. San Jose, CA.

23. Kraut, R.E., Miller, M.D., & Siegel, J. (1996). Collaboration in performance of physical task effects on outcome and communication. In: *Proceedings of CSCW '96.* Cambridge, MA.

24. Kraut, R.E., Gergle, D.,& Fussell, S.R. (2002). The use of visual information in shared visual spaces: informing the development of virtual co-presence. In: *Proceedings of CSCW'02*, 31-40.

25. Kraut, R.E, Fussell, S.R., &Siegel,J.(2003). Visual information as a conversational resource in collaborative physical tasks. In: *Proceedings of Human-Computer Interaction*,18, 13-49.

26. Kravets, R.,& Krishnan, P. (1998). Power management techniques for mobile communication. In:*Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM’98)*,157-168.

27. Kremer, U., Hicks, J., Rehg, J. (2001). A compilation framework for power and energy management on mobile computers.In: *Proceedings of the 14th International conference on Languages and compilers for parallel computing.* Berlin.

28. Kjaergaard, M., &Blunck, H. (2012). Unsupervised power profiling for mobile devices. In A. Puiatti & T. Gu (Eds), Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering: Vol. 104. Mobile and Ubiquitous Systems: Computing, Networking, and Services, 138-149. Berlin, Germany: Springer-Verlag. doi: 10.1007/978-3-642-30973-1_12.

29. Kuzuoka, H. (1992) Spatial workspace collaboration: a sharedview video support system for remote collaboration capability. In: *Proceedings of CHI ’92*, 533-540.

30. Kuzuoka, H., Kosuge, T., & Tanaka, M. (1994). GestureCam: a video communication system for sympathetic remote collaboration. In: *Proceedings of CSCW 1994*, 35-43.

31. Linjama, J. & Kaaresoja, T. (2004). Novel, minimalist haptic gesture interaction for mobile devices. In: *Proceedings of NordiCHI 2004,* ACM Press, 457-458.

32. Luk, J., Pasquero, J., Little, S., MacLean, K., Levesque, V., Hayward, V. (2006). A role for haptics in mobile interaction: initial design ssing a handheld tactile display prototype. In:*Proceedings of ACM Conference on Human Factors in Computing Systems (CHI ’06).* Montreal, Canada. 171-180.

33. Nygren, T.E. (1991). Psychometric properties of subjective workload measurement techniques: implications for their use in the assessment of perceived mental workload. In: *Human Factors*, 33, 17-33.

34. MacLean, K.E., & Enriquez, M. (2003). Perceptual design of haptic icons. In: *Proceedings of EuroHaptics 2003,* IEEE.

35. Orr, J. (1996). Talking about machines. Ithaca, NY: Cornell University Press.

36. Ou, J., Chen, X., Fussell, S.R., Yang, J. (2003). DOVE: drawing over video Environment, In: *Proceedings of MM 2003,* ACM Press, 100-101.

37. Ou, J., Fussell, S.R., Chen, X., Setlock, L.D., &Yang, J. (2003). Gestural communication over video stream: supporting multimodal interaction for remote collaborative physical tasks. In: *International Conference on Multimodal Interfaces,* ACM Press, Vancouver. 242-249.

38. Oulasvirta, A., Tamminen, S., Roto, V.,& Kuorelahti, J. (2005). Interaction in 4-second burst: the fragmented nature of attentional resources in mobile HCI. In: *Proceedings of the CHI 2005*. Portland, OR, AMC Press, 919 - 928.

39. Pirhonen, A., Brewster, S., & Holguin, C. (2002) Audio metaphors as a mean of control for mobile device. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems.*

40. Rao, R., Vrudhula, S., &Rakhmatov, D., (2003). Battery modeling for energy aware system design. In:*Computer*, 36(12), 77-87.

41. Satyanarayanan, M. (1996) Fundamental challenges in mobile computing. In: *Proceedings of PODC '96.* AMC Press, Philadelphia, PA.

42. Schuster, M. (2010) Speech recognition for mobile devices at google, In: *Proceedings of the 11th Pacific Rim International Conference on Trends in Artificial Intelligence*.

43. Smailagic, A. (1998) An evaluation of audio – centric cmu wearable computers. In: *ACM J. Special Topics in Mobile Networking*, 6(2), 59-68.

44. Stevenson, D., Li, J., Smith, J., &Hutchins, M. (2008). A collaborative guidance case study. In: *AUIC2008*, Wollongong, NSW, Australia. 33-42.

45. Wickens, C.D., & Hollands, J.G. (2000). *Engineering psychology and human performance* (3rd Ed.), Upper Saddle River, NJ: Prentice-Hall.

46. Zaykovskiy, D. (2006) Survey of the speech recognition techniques for mobile device. In: *Proceedings of SPECOM,*St Petersburg, 88-93.