NTNU – Trondheim
Norwegian University of
Science and Technology

A Mobile Decision Support System for High Risk Environments

Built Using Heads Up Displays and Intelligent User Interfaces

Gjøran Moen

Master of Science in Computer Science
Submission date: June 2014
Supervisor: Asbjørn Thomassen, IDI

Norwegian University of Science and Technology
Department of Computer and Information Science

Gjøran Moen

A System For Mobile Decision Support In High Risk Environments Using Heads Up Displays and Intelligent User Interfaces

Master Thesis, spring 2014

Faculty of Information Technology, Mathematics and Electrical Engineering
Department of Computer and Information Science

Problem Description

A system for decision support at the individual and organizational level, intended for military and emergency services use, shall be designed. The system should be able to contribute to the users' situation awareness and ability to solve given tasks cooperatively. This is done by considering the context, such as the user's spatial context and internal mental and physiological states.

The system shall be based on agent-based intelligent user interfaces and heads up displays in the form of wearable COTS hardware such as the Recon Instruments HUD. Architecture and subsystems shall be supported by research in relevant disciplines.

Abstract

This thesis describes the design of a system for decision support on the individual and organizational level, intended for use in military and emergency agencies. The system is intended to contribute to the users' situational awareness and ability to solve given tasks cooperatively. This is achieved by considering the context, such as the user's spatial context and internal mental and physiological states. The system is based on an intelligent agent based user interface and heads up displays (HUD) in the form of wearable off the shelf hardware such as Recon Instruments HUD products. Research in relevant disciplines covering, among other things, themes like physiological responses to psychological states and limitations of human visual perception is used to support design decisions.

Norwegian Abstract

Denne masteroppgaven beskriver et design av et system for beslutningsstøtte på enkeltmanns- og organisasjonsnivå, tiltenkt militær- og nød-etater. Systemet er tiltenkt å kunne bidra til brukernes situasjonsbevissthet og evne til å sammen løse oppdrag. Dette oppnås ved å ta hensyn til kontekst, slik som brukers spatiale kontekst og interne mentale og fysiologiske tilstander. Systemet bygger på et agentbasert intelligent brukergrensesnitt og heads up displays(HUD) i form av bærbar(wearable) hyllevare som feks Recon Instruments HUD. Forskning i relevante fagområder om for eksempel fysiologiske responser på psykologiske tilstander og begrensninger i den menneskelige synssans er brukt til å understøtte designavgjørelser.

Preface

This thesis is submitted to the Norwegian University of Science and Technology as part of the requirements for attaining the degree Master of Science. The thesis has been written as the culmination of Master Studies at the Department of Computer and Information Science of NTNU, Trondheim.

Acknowledgements

First I wish to thank my advisor, Asbjørn Thomassen at the Department of Computer and Information Science of NTNU, for great guidance and for allowing and encouraging me to go with my own idea for my Master Thesis. Next thanks to friends and family for moral support throughout the final stage of my Masters education. Finally to my house mate Marte for allowing me to use her living room as my de facto workspace for the semester.

Gjøran Moen
Trondheim, June 7, 2014

Contents

List of Figures

Chapter 1

Introduction and Overview

1.1 Background and Motivation

The idea for this thesis originally came from the authors experience with military service and from experiencing the limitations and possibilities in the communications technology of the defence sector. There are several C4I capabilities not being fully utilised in military settings, even though they have seen widespread use and adaptation in the civilian world. This was supported in a small survey done among serving military personnel in [Moen, 2014]. Even old tried and true paradigms such as Short Messaging Systems are not utilized much, even though the capabilities are very much in place. It is the authors belief that this is not due to these capabilities being unsuited for military adaptation, but rather a result of outdated user interface design principles applied to the C4I technologies in active use today. This thesis therefore tries to demonstrate a system design that would enable the users of these C4I technologies to make use of their full potential. Given the trend of reduced defence budgets throughout the western world, the economical cost is also an important factor in any system that aims towards military use. Given the stressful environment military users often operate within, it would make sense if this necessarily mobile system would provide them with a natural information flow that would contribute to their situational awareness.

1.2 Inspiration

Besides personal experience there are many sources of inspiration for this type of project. Several projects both under development and in active use exists, demonstrating both the need and demand for systems similar to the one outlined in this thesis. Rheinmetall Defence has seen wide coverage of their Gladius system after they transferred it to the German Bundeswehr in 2013. The system is planned to be deployed with German forces in Afghanistan during 2014.[RheinmetallAG, 2013] Consisting of several subsystems such as clothing, load carrying equipment etc, Gladius also incorporates a state of the art heads up display. The HUD includes both night vision and augmented reality capabilities, enabling access to functionality and information such as maps, waypoints and tracking of both friends and foes. The system has been described in several online articles, such as this one, [Bittel, 2013]. In this article, Bittel starts out from the get go comparing the system to modern computer games such as Halo and Call of Duty. Whether this was part of Rheinmetall's inspiration is unknown. Still, the idea of pulling inspiration from the gaming industry, which has spent the last couple of decades fine tuning the HUD interfaces for their games, does not seem like a bad one. The US canceled and then restarted in 2008 a similar system called Land Warrior. The latest innovation for this project is the Battlskin helmet from Revision. [Macias, 2013] It also incorporates heads up display capabilities in order to enable situational awareness aids. Many more such systems and projects has been publicly confirmed, such as the Canadian Integrated Soldier System Project[Pugliese, 2013] and the Personal Eye System by MIL Sistemika[Sistemika, 2014], the latter being android based and commercially available to anyone through the Google Play Store. So, the idea and need of a situational awareness and decision aid for the dismounted soldier clearly has wide acceptance in the defence industry. How far these projects have come in terms of research and specifically what functionality they have implemented is difficult to say, since most defence industry companies and military organisations naturally tend to limit how much of this information they reveal to the public. Still, we can clearly see the relevance of these types of systems through these public announcements.

1.3 Goals and Research Questions

The main goal of this thesis is therefore to clarify and demonstrate how a context sensitive situational awareness system for use with infantry soldiers should be designed and what functionality it can be expected to provide. In addition, identifying challenges and opportunities in interfacing with a user that are specific to a high risk environment. With this background the following research questions are posed:

- What human factors needs to be considered when interfacing with a human user operating in high risk situations?

- How can a pervasive system in a high risk setting recognize and classify context? Specifically:

  - Detect critical situations
  - Detect and consider implications of user emotions

- How can context be used to inform reasoning processes?
- How can these reasoning processes and their results be used to create an increased situational awareness within the user?

1.4 Research Method

The focus of this thesis has been to establish a interdisciplinary basis and clear guidelines on how such a system should be designed in order to determine its feasibility. Therefore mainly theoretical research, largely into disciplines outside the realm of computer science, has been the focus. The findings from this research was then used to warrant the subsequent design decisions. A design based on these decisions is then presented, intended as a proof of concept for the idea.

1.5 Thesis Structure

The goal of this thesis is to present the idea of a system and document its feasibility. Therefore some theoretical background is presented in the next chapter. Thereafter this basis is used in supporting the design decisions made. Then a set of system models resulting from these decisions are shown before the models and decisions made are summarized and discussed.

1.6 Military specific terms

Since the described systems intended use environment is in a military setting, some military specific terms will be used in describing system usage scenarios. For clarity, a list of such terms is provided here.

”Blue on Blue” Term used to describe a situation where some unit unintentionally engages another friendly unit.

BLUFOR Term used for all friendly forces in a given operation.

OPFOR ”Opposing Force”, the forces opposing the BLUFOR in a given operation.

TIC ”Troops in Contact”. Term for a situation where a unit has made contact with enemy elements. The result may be a fire fight or some other extra high risk situation.

CBRN Chemical, Biological, Radioactive, Nuclear. Collective term for environmental hazards, both naturally occurring as well as intentionally deployed agents.

C4I systems Short for Command, Control, Communications, Computing and Information Systems. Umbrella term used for all military specific systems in the C4I categories.

BMS Battlefield Management System. A mobile communication and information management system that aids military commanders when managing forces, resources and information across an operational area, such as a battlefield. The BMS nodes communicate over a Tactical net.

TacNet Tactical Net. Data transfer enabled network implemented through mobile communication technologies such as radios and relay stations. The Tactical net is independent of any stationary communications infrastructure, but may make use of it when available.

Chapter 2

Theory and Background

This chapter will walk through some relevant theory in order to provide a background for the design decisions to be rationalized in the following chapter.

2.1 Early work on agent based intelligent user interfaces

The Intelligent User Interfaces(IUI) paradigm is relatively new and has therefore not yet generated the same level of standardized guidelines of design as for example the traditional Graphical User Interface(GUI). IUI's brings both new opportunities as well as challenges for the UI designer that needs to be considered. In [Maybury, 1998] the need for IUI's was described as a result of the ever increasing availability of information on both local and global networks. This availability has made the task of disseminating the available information an impossible one for any human being. Therefore the need for more efficient, effective and natural interfaces that are able to present these huge amounts of data in a meaningful way. Since then the availability of data, and subsequently the need to interpret it, has increased immensely. Now that this flow of information has been channelled into our pockets through mobile technology, we have a constant need of data sense making services. Also, the wearable technology, described in section 2.2, is gaining ground, especially in military and emergency agencies and in the world of extreme sports. In these environments time and situational awareness are critical, even to the extent that lives depend on it. It is here that IUI's can provide the extra edge by helping the user in interpreting all the available data in an efficient and effective way. By delegating menial and data heavy tasks to the IUI the user can avoid falling victim to information overload and keep his mind at the task at hand. Another important aspect of IUI's is the reasoning process that allows these interfaces to provide informations services to the user. In [Ishizaki, 1996] the process of multi agent design(MaDes) is described and exemplified. This process utilizes so called design agents. These are agents that control certain elements of the UI layout and designs and redesigns these elements on the fly in order to provide an interface design. One example is the use of attention actions by the design agent. By using graphical cues the design agent can bring the users attention towards a piece of information that is relevant to the user in the given context. Other research on agent based information dissemination includes [Bayardo Jr et al., 1997] and [Maes et al., 1994], that further understates the usefulness of intelligent agents in information rich environments. In order to provide a natural intelligent user interface new interaction modalities has and should be explored further. In [Cohen et al., 1997] a system is described that uses multiple input modalities to enable the IUI to infer the intention of the user. Since voice interaction can be prone to create misunderstandings, multiple modalities are combined so that the system can more accurately interpret the users intention. Trough the Quickset interaction system they, among other uses, enable controllers of a military distributed simulation system to effectively control the simulated environment through a natural interface. In one application, voice and pen drawing modalities are combined in order to build simulated military scenarios. Similar to the MaDes systems, the Quickset system is also agent based. One of the agents described in the paper is the Application Bridge Agent. This agent acts as an interface between the underlying application and the Quickset interface system. In other words, Quickset can be adapted to a wide range of systems simply by rewriting the Application Bridge Agent. Next we will look at some new paradigms and technologies that is making IUI's more and more relevant, and feasible.

2.2 Wearable Computing

What does one expect from a wearable computer system? In [Starner, 2001a] the following definition of wearable computing is presented: "provide portability during operation; enable hands- free or hands-limited use; can attract the users attention, even when not in active use; can run continuously; and attempt to sense the users current context"

The author from this tries to define and describe an ideal wearable computer. He lists 4 ideal attributes.

- Persist and provide constant access to information services.
- Sense and model context.
- Adapt interaction modalities based on the users context.
- Augment and mediate interactions with the users environment.

These points illustrate a vision of wearable computing that correlates very well with what we are trying to do with our system. However, according to Starner, there are several challenges to wearable computing that needs addressing.

The primary challenge of wearable computing is power usage. Given the fact that a wearable computer by definition must go where its user goes, it cannot rely on any constant source of power. In order to limit intrusiveness of the wearable system, large, heavy batteries are also out of the question. Starner also points out several political and ethical challenges when it comes to some power supply solutions. Steps must therefore be taken to limit the amount processing needed and thus power requirements for the tasks performed by the system. The challenge of power consumption also leads to the issue of heat. These two issues together shows the need for specially designed processors for use in wearable computers, focusing in MIPS per watt, more than maximising processing power.

Further, in his part two on challenges of wearable computing[Starner, 2001b], Starner describes the challenges surrounding networking. In this issue also, power consumption is a constraint that needs to be taken into account, even when communicating between on body devices. Limited service availability and interoperability are other constraints that come into play. Wearable computing requires a wide range of wireless networking capabilities, on body, near body and WAN communications. This means one should try to enable these capabilities on a single, or as few as possible radio devices. To do this new strides must be made on interoperability between networks of different scales and ranges. This and other measures should be taken in order to maximize bits per second per watt, a measure Starner deems much more appropriate than bps in the wearable setting. Also, some services may not be available at all in some remote areas. One suggested solution to this problem is the use of mobile repeaters mounted in vehicles or other crafts that a user usually brings with them into these areas. This approach brings out a whole new set of challenges, such as service standardization, synchronization, security etc. A solution where intelligent agents, with the help of caching and revision control, can emulate remote access until such access is restored is also suggested. Privacy is also stated as a key challenge for wearable computing, which is obvious when viewed in a civilian setting. Starner also states that privacy is not the same as, but actually often comes in conflict with the issue of security. These are issues that would be balanced very differently in a military setting versus the civilian setting of emergency response. Next issue is interface design. Here Starner illustrates a very interesting idea. A kind of modular interface where user interaction hardware is made up of different modules. Each module represents a certain modality, be it video, audio or text input. Using the on body network to communicate this enables the user to pick and choose sensors and modalities on a current need basis. The ”A day in the Life” case written in [Starner, 2001b] illustrates the idea rather well. The issues of GUI and Intelligent User Interface will be examined in later sections.

If the described challenges can be tackled, there are a number of possibilities that are enabled through a wearable system. We have already seen that modularity of modalities can be achieved, enabling the user to tailor the system capabilities to his needs, on the fly. Intellectual tools, as Starner puts it, can help augment the users ability to perform cognitive tasks through psychophysical analysis of the user, pure and simple computing power or through intuitive data presentation techniques. By coupling these intellectual tools with communications capabilities, we can use wearable systems as a means to facilitate better cooperation and team work.

The idea of using wearable technology to build collaborative systems is far from a new one. In [Kortuem et al., 1999] the authors describe a collaborative wearable system for use in manufacturing or maintenance teams where the technicians tend to be very mobile. The goal was to tackle inefficiency related to the communication technology and processes currently in use at computer maintenance centres. They find that limitations of pure voice based communications such as mobile phones makes them unsuited for communicating complex expert information. Also, the keeping and updating of maintenance logs are complicated by lack of documentation, while the completion of practical tasks are hindered by cumbersome and impractical communications equipment. This is where wearable technology helps alleviate these problems by enabling a natural way for these collaborators to communicate remotely. First person video feeds enables expert advisers to see what the on site technician can see, making the advising process much more effective. Hands free operation enables the technician to work and receive expert guidance at the same time. Also, the data transfer capability enables constant and up to date data log access.

All these possibilities brought by wearable computing enables us to build pervasive and ubiquitous systems, described in the next section, that are always physically with the user and can tend to his information needs at any time.

2.3 Ambient Intelligence

In the literature we often see the terms ”Pervasive Computing”, ”Ubiquitous Computing”(UbiComp) and ”Ambient Intelligence”(AmI) being used interchangeably. Often to describe this mysterious omnipresent system that knows and sees its surroundings and then acts upon them. It therefore seems prudent to clarify these terms and their relations.

The term Ubiquitous Computing was coined by Mark Weiser in 1991[Weiser, 1991], and refers to computing technology that disappears from the users awareness by becoming an integral part of every day life. According to [Singh et al., 2006] the terms UbiComp and pervasive computing are often used interchangeably in the literature, but they state that they are conceptually different. Pervasive systems can be said to build on the earlier mobile and wearable computing systems, incorporating context awareness as a the key capability moving it past these earlier systems. Given the cost of multiple complex and costly specialized devices needed for implementing a ubiquitous system based solely on mobile technology, pervasive computing seems to be the most viable next step towards Weisers vision. We can draw from this the distinction that pervasive computing is not the same as UbiComp, but rather an important building block thereof. The next term we need to tackle is "Ambient Intelligence". [Aarts and Wichert, 2009] puts it like this: "Ambient Intelligence (AmI) is about sensitive, adaptive electronic environments that respond to the actions of persons and objects and cater for their needs." In other words, we can view the ambient intelligence as all electronic sources of information available in a given context. Examples could be the signals from a GPS satellite, WiFi enabled sensors or mobile networks etc.

Thus we can describe the relationship between these concepts as follows: Pervasive systems, operating in an environment of ambient intelligence, can work together in order to form a ubiquitous computing system. The system described in this thesis belongs to the pervasive category, operating within the frameworks of AmI and UbiComp.

2.4 Explanations

The concept of explanations has grown parallel to the emergence of expert systems. The purpose of explanations has primarily been to build trust and acceptance in the users, enabling the expert systems widespread acceptance. In the process several methods for explaining AI reasoning processes has been developed. Reviews of explanation methods and principles for heuristic, Bayesian and case based reasoning can be found in [Lacave and Diez, 2004], [Lacave and Diez, 2004] and [Sørmo et al., 2005] respectively. One of the earliest and simplest methods, described in [Lacave and Diez, 2004], is canned text. Here one simply uses a text template and fills in the blank data points with values from the reasoning process. The message template shown in section 3.13 is a perfect example of how this method could be used.

2.5 Augmenting reality

As shown in the introduction there are many similar systems both under development and in active use. Many of these systems uses the relatively new Augmented Reality(AR) technologies. AR is typically based on a "see through" hardware platform, such as a smartphone with camera or google glass. By making use of the hardware camera, display, sensor and processing capabilities, AR applications aims to draw 3D objects on top of the world view as if they were a part of that world. This article aims to demonstrate a system based on a non AR LCD screen, and a quick look on both methods strengths and weaknesses is therefore necessary. In [Kaufmann and Dünser, 2007] an educational AR based system is evaluated. The system uses an AR capable Head Mounted Display to enable students to create 3D geometric models. According to the evalutaions results the AR system achieves a significantly higher usability rating than a traditional desktop based system built for the same task. However, the AR system scores somewhat low when it comes to conforming to users expectations. They do not however, go into details on what form of expectations the system struggled to adhere to, or how one might remedy these limitations. They do claim to improve this score over the next development iterations, but still without much detail. It seems research done on how AR systems affects the users expectations is limited. What would be interesting to see is research on what level of data accuracy users would expect from an AR interface, compared to a more traditional LCD/HUD based interface. Further, [Kaufmann and Dünser, 2007] shows that users of the AR system tended to experience so called "Simulator Sickness" after 20-60 minutes of using it. This condition is similar to motion sickness, and the evaluation results shows eye strain and tiredness to be the most common symptoms experienced. In a recent article in the Norwegian "Teknisk Ukeblad" magazine, about an AR system for armoured vehicle operators, this is further corroborated by a statement by an officer involved in the project. He states that wearing the AR goggles, currently they are using the commercially available Oculus Rift, over time causes fatigue, thereby limiting their continuous use in operations. [Urke, 2014] Generating AR imagery, with 3D objects drawn as an integrated part of a real world image, is computationally expensive, and thus puts a further strain on the already limited system resources typically found in wearable devices. A standard LCD based HUD, like the Recon MODLive HUD, on the other hand, has the limitations of not having a see through capability[1], and cannot be in the users focus or attention at all times. Also, the 2D nature and limited real estate of the display creates extra challenges for the design of a user interface for the system.

[1] Newer models released by Recon Instruments has the potential for AR operation as a result of added sensors, camera and increased processing power

2.6 Designing intelligent user interfaces

[Maybury, 1998] lists a set of benefits that IUI's promises to provide:

- Comprehension of complex data
- Generation of natural and meaningful presentations of data
- Automatic performance of delegated tasks
- Management of the user interaction and interface

How should one proceed in order to design an IUI that provides these benefits? According to [Ali et al., 2012] the field of IUI's is so new that no standardized guidelines exists. While no model or guidelines for IUI's exists, he suggests 5 steps that should be followed in order to design good Intelligent interfaces:

Step 1 Understanding the problem space

Step 2 Analyse and build a set of alternative prototype designs

Step 3 Develop working models and evaluate them

Step 4 Remove drawbacks and incorporate suggestions

Step 5 Repeat 2-4 until done.

In other words a very general approach. In order to have some sort of starting point for a good design it makes sense to use architectures from earlier similar systems as a basis. This goes under step 2 above, where prototype designs are suggested. One approach that has seen widespread use with IUI's is the use of intelligent agents. In [Heinroth et al., 2011] the Activity Sphere Architecture(ASA) is presented.

An activity sphere(AS) is described as a bubble around a current task, limiting data flow and actions to that which is relevant to the current task. An example AS used in the article is "cooking", another more relevant to the context of this thesis might be "reconnaissance patrol". The ASA revolves around a Sphere Manager. The task agent subscribes to event notifications from the manager and returns an action as a response. The ontology manager manages the ontological knowledge of the sphere. It also sports an Interaction agent that has responsibility for managing interaction and interface as well as handling multi modal inputs from the user. Last we have the planning agent. It is responsible for the goal tracking and overall strategy of the system.

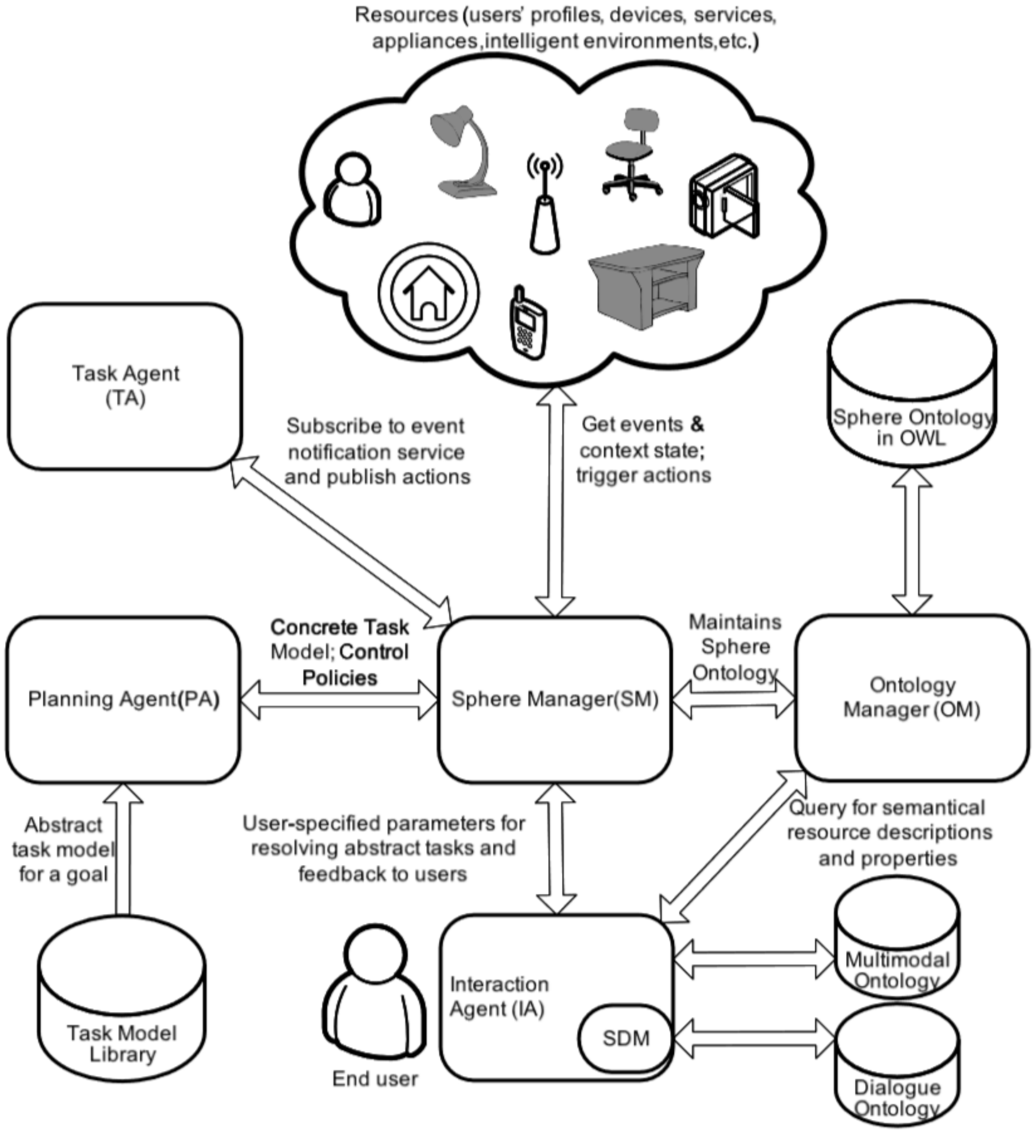

Figure 2.1: The Activity Sphere Architecture. Taken from [Heinroth et al., 2011], fig 1.

[Heinroth et al., 2011] focuses on the Interaction Agent. There is a set of issues that this agent needs to address given an adaptive multi modal interaction approach.

The first issue is allocation. This has to do with the choice in modalities to be used in the interaction. What method is best suited for performing user input? Voice-, text-, tactile- or otherwise based input modalities are each better suited in some given contexts. Tactile input would be impractical in a context where the user is using both hands on some other task for example.

Next issue is instantiation. This means selecting suitable parameters for the allocated interface modality. A simple example might be sensitivity of a computer mouse or the activation threshold of a voice detection device.

Evolution is the issue of changing parameters of the interface during interaction with the user. This might mean managing a dialogue of negotiation between the user and the system.

Another approach is to use adaptive interfacing. In [Arai et al., 1993] recursive fuzzy logic is applied in order to achieve a gradual adaptation to the changing user state. By using recursiveness in the logic, sudden changes in perceived user state does not result in a unpredictable interface experience. This approach seems particularly useful when dealing with complex states such as human emotions and the unpredictable sensor data from physiological sensors.

2.7 Physiological and psychological effects of emotional stress

How does stressful emotions such as fear and anger affect the human minds ability to process information? Can we measure the level of emotional stress through simple sensors? These questions will be important when designing an interface with the aim of communicating efficiently with a user that may or may not be under the effect of emotional stress. The need for us to ask the second question lies in the answer to the first. What are the effects of emotional stress on humans ability to handle information flows? [Christianson, 1992] reviews several studies made on the effects of emotional stress on the efficiency of human memory. They conclude that emotional arousal do have a clear effect on memory, but not in the way illustrated by the Yerke-Dodson inverted-U law[2] and Easterbrook hypothesis, which seems to have been the prevailing theories for several decades. The author concludes that there is no evidence that suggests such a simple relationship between emotional arousal and the efficiency of human memory. The author does however mention studies that indicates that changes in attention patterns can represent part of the reason why emotional events are remembered differently than ordinary every day events. From earlier studies they infer that emotional stimulus leads to heightened attention towards details of the emotional event. This in turn leads to impaired attention towards peripheral details. Another study, [Keinan, 1987], determines that the decision making process is affected when the decision

[2]The Yerke-Dodson(1908) law states that there exists a one dimensional relation between level of emotional arousal and memory efficiency represented by an inverted-U curve. See Figure 1 in [Christianson, 1992]

maker is put under emotional stress. The study states that decisions made under emotional stimuli tended to be made before all alternatives were considered. Also, the subjects tended to review their alternatives in a non systematic fashion. The author describes one possible explanation being that attention put towards a perceived threat takes from the capacity to handle the current task. From this we can generalize that the main psychological effect from emotional stress is the "hijacked" attention of the user, limiting his ability to process details not related to the object of emotional stimuli and potentially taking his attention away from the task at hand. Next question is then, how can we detect the emotional states that triggers a hijacked attention? Are there any measurable physiological effects that can help us determine if these stimuli are effecting the user? In [Rainville et al., 2006] the authors shows results from a study of 43 subjects that are consistent with the idea that distinct patterns in cardiorespiratory activity can be associated with distinct emotions. They even provide a neat heuristic decision tree for determining emotional state based on cardiorespiratory activity. This means that, according to this study, sensors for measuring heart and respiratory rates would enable us to classify a users emotional state.

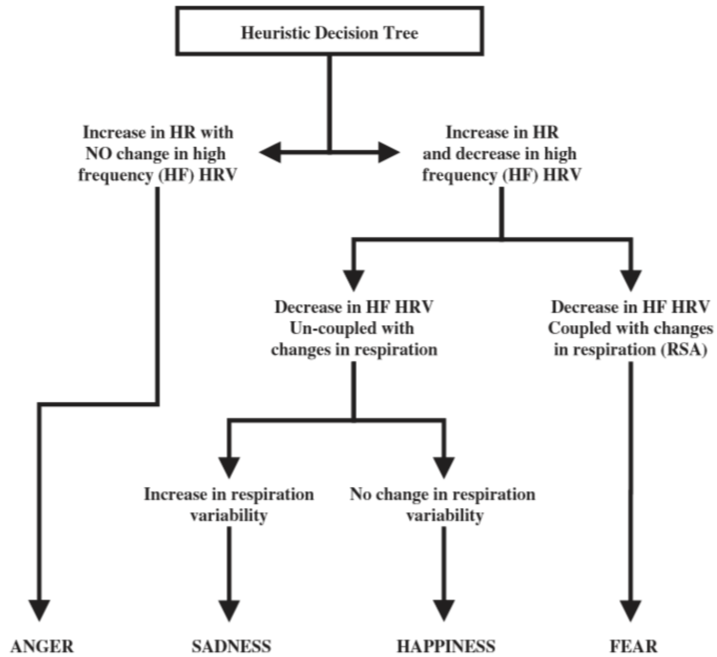

Figure 2.2: Heuristic Learning Tree for determining emotional state. Taken from [Rainville et al., 2006]

2.8 Human Vision

The human eye is a complex and impressive organ. However, it also has its shortcomings, that we must take into account when designing a system that mainly communicates with the user through his vision. Human color vision is based in three classes of ”cones” that are distributed in a ”mosaic like” pattern, but not necessarily evenly so, as we can from the images shown in [Roorda and Williams, 1999]. A good overview of how the eye works and how it is constructed can be read in section 5.7 of [Hecht, 2002], however, I will make a short summary of the key points here. The sensory cells that enables the human eye to register light and color comes in two main types, rods and cones. The cones are sensitive to colors, whereas rods are not. Cones can be further subdivided into three classes, red, green and blue.[3] These are sensitive to light of the colors red, green and blue respectively. Most of the cones are located in the fovea, the central part of the retina, most densely packed closest to the fovea centre. This means that only a small part of the eyes field of view provides the most detailed visual representation. One of these cone classes, the blue class, constitutes only a small percentage of the total number of cones as seen in samples from [Roorda and Williams, 1999]. Roorda also argues that the somewhat uniform distribution of blue cones are disrupted by the density of cones in the fovea, therefore limiting the number of blue cones present there. Further away from the fovea the density of cones decreases relative to the increasing density of rods. The rods are more sensitive than the less responsive color cones and provides the eyes peripheral field of view. [Ancman, 1991] studies the limitations of color vision the peripheral field of view by displaying different colors on CRT monitors placed in the peripheral view of the test subject. The author finds that colors can be misrepresented in the peripheral vision and the subjects, as a result, could not identify the displayed colors correctly. The color found to be identified correctly the furthest away from the fovea was blue, at a maximum of 83.1°. Blue was also found to be wrongly identified only 5% of the time compared to 63 and 62% for red and green respectively. She also found that the emotional state of subjects has some effect on the width of their peripheral vision, in accordance to the inverted-U rule described in 2.7. As shown in the previous section this is probably an overly simplistic model for these effects, but it nonetheless illustrates how stressful situations affects the information processing and other abilities of the user.

[3] I denote these cone classes by color names for simplicity, as the long-, medium-, and short wavelength sensitivity denominations makes less sense in this context.

2.9 Context and Situational Awareness

For a system to be able to provide context sensitive information, with the goal of improving the situational awareness of its user, the system must first achieve context awareness for itself. We therefore need to define context and how to sense it. In his PhD Thesis[Kofod-Petersen, 2007] Kofod-Petersen looks at context through the eyes of human reasoning. He describes context as a situation assessment owned by a sensing entity. In other words, context is a subjective construct. This human view on context can be transferred to an intelligent sensing system. This means that a context aware system needs the ability to sense and determine context from its own point of view. Determining context is not a trivial problem, as it relies heavily on how we define and quantify knowledge of our environment. Kofod-Petersen points to [Ekbia and Maguitman, 2001] who claims that context has not been given much attention by the Artificial Intelligence community. They have therefore turned to the philosopher John Dewey in order to categorize context. This results in two main categories of context, spatio context and temporal context, both constituting sub parts of a complete context. As a simplification we can view these parts as external and internal context respectively. Externally we have the observable state of the environment. The internal context can be exemplified by the sensing entity's previous knowledge of the world and of itself. When looking at both internal as well as external as part of the same greater context, Kofod-Petersen shows that one has to move away from the traditional AI view of separating the reasoning process from the world which the entity inhabits. His next step is the notion of a action oriented intelligence where context is used as a tool for selecting the best course of action.

2.10 Challenges of a Military Setting

When operating a wearable system in a military or emergency setting, there are a whole range of limitations and operational aspects that needs to be taken into account. Here's a short list of a few such considerations that needs to be made.

Operational Security Keeping operational data from getting into the wrong hands is more critical than in an average civilian setting.

Vision The human vision is ground personnels most important tool for analysing a situation, making this resource both scarce and exceedingly important to their ability to do their job.

Speech and hearing These are a teams most important channel of communication, both face to face and via radio. This means that both hearing and speech are already overloaded resources.

Life and death decisions Decisions made in error can have dire consequences, possibly endangering allied or innocent lives.

Chapter 3

Design Decisions

In her contribution to the Computer interaction handbook, [Watzman, 2003], Suzanne Watzman says on usable interface design: "When we succeed our products become effortless, even pleasurable to use". This should be the goal of our design process, to make the issued communications equipment effortless to use for the soldier who depends on it. Watzman also suggests 5 criteria for good design that should work as a good guideline in any interface design process.

Appropriate Is the design solution fitting to the intended usage environment and users?

Durable Can the design last, is it adaptable to evolving usage environments?

Verifiable Is the design tested and evaluated in the usage environment?

Impact Does the product resonate with the user making it desirable and therefore adoptable?

Cost Effective Does it have low development and maintenance costs?

This chapter will try to realize a design according to these criteria and others described in earlier chapters.

3.1 Intelligent user interface

In the described setting of our system, there is little room for extensive hand gesticulation or touch interfaces, given the use of both hands in most tasks delegated to the user. Add to that the fact that gloves are usually worn and in an often chaotic environment, and we see clearly the need for a simple and effective user-system interface. These requirements however puts some constraints on what modalities we can use for user direct input. As a result the system will need to compensate for this lack of user input with some degree of autonomy, limiting the amount of choices proposed to the user at any one time. The discussed Intelligent User Interface paradigm described in 2.1 will provide a suitable framework for enabling this automation. Drawing information from multiple modalities, such as sensors connected to the system, the IUI can adapt to the users context and provide semi automatic dissemination of the data flow made available to him, freeing both his hands and brainpower for other, more critical, tasks. In section 2.6 a process of designing a IUI is described. This, and the next chapter can be seen as step 2 in the design process, where we analyse and develop a prototype design.

3.2 Agent based dynamic design

In [Moen, 2014] the [Ishizaki, 1996] paper was used to argue for an agent based architecture, emphasizing flexibility through modularity. Also the idea of the design agent was pulled from the MaDes paper. In section 2.1 the IUI approach and the use of intelligent agents in the management of this new type of user interface is further described. The use of these intelligent agents will provide the system with flexibility and scalability through the modular nature of agent based system. This will allow the system to be scaled and configured to match a wide set of operational environments and requirements. This allows the balancing of available functionality with the computing costs and the power consumption that goes with it.

3.3 Hardware Components and Modularity

The vision of this system is that of a modular one, that can be adapted to a wide variety of tasks and usage scenarios. For economical reasons, both in development and use, it should be largely based on commercially available off the shelf hardware(COTS). For this thesis the Recon HUD by Recon Instruments has been selected as the intended platform, allthough the design principles can also be applied to later models from Recon Instruments or other similar COTS products. Section 2.5 showed that a LCD HUD can be just as relevant as the new and more popular AR technology. The reduced cost and less required processing power of using a traditional display is key in keeping this relevance, especially in a wearable setting. Through either a Bluetooth or cable based on body network the HUD can connect to different devices carried by the user. What devices this might be is discussed bellow. A cable based on body network would have the added benefit of being able to distribute power as well as data, at the cost of being more cum-

Figure 3.1: Ballistic helmet with HUD goggles, sensor module and bluetooth remote. Taken from [Moen, 2014].

bersome to dress in and out of. In addition to the HUD unit, the base system is dependent on a hardware interface connecting the on body network to the data capable long range communication device employed by the user. The principle of modularity is applied to both software and hardware in order to end up with a truly versatile system. The HUD unit can expand its capabilities by adding devices to its on body network. Some of these rely on the implementation of software modules, typically intelligent agents, that can manage these added resources and integrate them into the systems mode of operation. The components can include but is not limited to:

Sensor Module Expands sensor capabilities of the HUD unit. The sensor module itself is modular, enabling plug and play expandability. Examples of sensors: Ambient sound, rain, heart rate, CBRN etc. This is the only hardware component that might require some fabrication efforts.

Digital camera Enables image processing capabilities such as facial recognition as well as simple photography and video streaming.

Smart phone, tablet or laptop For touch screen or keyboard interaction capabilities, extra processing power and other typical capabilities for these devices.

Headset For voice input and sound output. Must also enable use of the ordinary push-to-talk voice communication mode.

Storage Extended digital storage capabilities. For example for images from a camera. Could also compensate for unreliable data connection by storing more data locally.

Power source Extended duration of operation and potential for more processing power.

3.4 Software Framework

The planned hardware platform is Recon Instruments which runs the Android operating system. The application must therefore be built on the Android API. Since the Android API is Java based, we can run the Jade, Java Agent DEvelopment framework, on it. Jade is a framework for running agents, based on Java, over a set of distributed agent containers. The Jade developers has shown this to be feasible through the demo app presented in their Android Tutorial. [Caire, 2012]. How and in what mode the Jade framework will be run on the mobile Android nodes is described in subsection 3.4.10 in [Moen, 2014].

3.5 Agent Categories

To simplify and clarify the design of the agent architecture a set of agent categories has been made. In [Moen, 2014] the categories ”Informers”, ”Service Providers” and ”Consumers” were used, but these were found to overlap too much to be of much use in clarifying the architecture. The first category is resource agents. They are each responsible for managing a given resource that are available to the system, as well as providing services relating to that resource. In the case of the Blackboard Agent, this includes updating the local knowledge base with new information coming in from other agents, removing outdated data and providing responses to data requests. The idea of defining resource agents as agents managing resources and making them available to the reasoning agents was taken from [Bayardo Jr et al., 1997], where resource agents are used to handle database and http resources.

Next category, specialist agents, consists of agents built to provide expert reasoning relating to available data. They are responsible for determining the current context within their frame of expertise. An example being the Mapping Agent that may identify an event where the user is approaching a known area of interest, such as a minefield or a dangerous chemical spill area. Next are the event handling agents. This category is principally meant to hold the Notification Agent, further described in 3.9 and 4.1, but may include a Task Progression Tracker and a Shared State Agent. The event handlers receives events in real time from the Specialist Agents. The handler then makes a decision on whether any immediate action is to be made from the event. Lastly we have the Services Agents. These guys provide a specific computing service to other agents. This might be a heavier task such as image processing calculations that might hinder the real time operations of certain agents. Bellow is a list of categories and the envisioned agents within them:

- Resource Agents
  - Sensor
  - Network
  - Blackboard
  - GUI
- Specialist Agents(responsible for reasoning process, classification of states etc)
  - Mapping
  - GunShot Detection
  - Emotional state
  - BLUFOR tracking
  - Imaging agent
- Event Handling Agents
  - Notification
  - Shared State
- Service Agents
  - Text Generation
  - Image Processing

3.6 Shared Knowledge Base

The knowledge base of the system is shared among all agents through the Blackboard agent. The Blackboard agent is tasked with servicing information requests as well as keeping the KB up to date. This way any read-write conflicts can be handled in a trivial way, such as first come first served.

The KB itself will consist of two main information categories. The first is ”classified event cases”. These are representations describing occurring events that have been identified and classified by the system. Together they represent the ”working memory” of the system or, in other words, its recent contextual history. These cases can be compared to the cases of the Case Based Reasoning paradigm, but differ in a key aspect. The history of cases is not necessarily used when generating new cases, as this task is performed through the expertise of the specialist agents triggered by sensory events. This contextual history may be used in the reasoning process of some agents, for instance in the event handling Notification Agent. Any agent subsequently working on a case can provide it with extra attributes. This means that the number of attributes a case can have varies with the number and types of agents populating the system. Each case is identified through a time stamp value, enabling agents to use these to build historical data structures such as Markov Chains for use in reasoning processes. In order not to fill up the working memory of the system, the timestamped cases must have a limited lifetime, after which they are ”forgotten” by the system. Any information that might be useful at a later time must therefore be stored in the systems long term memory.

The next category are meta data. The meta data has no limited lifetime and can thus be viewed as the systems long term memory. These are attributes that are continually updated and are put to continuous use by Specialist and other agents. Examples of attributes are current GPS coordinates, map artefacts, message logs etc.

The data processing is a continual and collaborative process done by the agents. An example of this collaborative effort can be shown for when the GPS sensor updates the latest GPS coordinates. This happens at most four times every second, or at a frequency of 4 hz. The sensor agent, which is responsible for gathering sensor data, then sends a data package containing the coordinates to the Blackboard agent. The Blackboard agent then generates a new case and a unique id, based on a timestamp, for that case. The case and ID is then distributed to all relevant agents for review. One of these agents is the mapping agent. When receiving the new case the mapping agent updates the ”lastGPSLoc” value in the mapping metadata. Next it adds mapping specific attributes of the case it received, such as distance to and type of the closest area of interest stored in its mapping metadata. In this case lets say these values are ”50m” and ”MINEFIELD” respectively. Next it ads a classifier to the case’s class array, ”ProxMF”, that classifies this case as a

”minefield proximity case”.[1] If the agent has nothing to add to the current case it is returned to sender as is, otherwise the updated case array is returned. Had the case received by the mapping agent not contained GPS coordinates, it might add the latest GPS coordinates saved to the KB.

3.7 Agent Interaction and Anatomy

In order for the agents to cooperate, a framework for interaction between them needs to be established. The data structures and anatomical properties required for these interactions must also be defined. Each agent will need a working memory, keeping a subset of the system knowledge needed to complete its tasks. The interaction requires that the agent knows what agents are operating in its environment and what services they provide. The agent therefore needs to maintain an address book containing this information. We call this data structure the ”yellow pages”. Next the agent needs to know what services it can provide under the current circumstances. To do this a data structure is kept that contains all potential services it can provide and the resources or third party services required to be able to provide that service. Lastly the agent needs to know what agents are currently subscribed to each service.

The interaction process itself is broadly divided into two phases. The first, setup phase, is where the agent builds up its yellow pages and communicates the services it has to offer to other agents. Next phase is the runtime phase, where the agent has built up its yellow pages and subscribed to relevant services. In this phase agents work cooperatively to provide the best service possible to the user.

3.8 Gunshot Detection Algorithm

A gunshot represents a clear escalation of the situation the user is in. We therefore wish to be able to classify gunshot events in order to adapt the user interface to this sudden change in context. Many algorithms exists for classifying signal patterns, such as those found in an audio signal. We wish to classify whether an audio signal represents a gunshot or not in order to determine the nature of the users context. In [Chacon-Rodriguez et al., 2011] a set of hardware implemented pre processing algorithms are evaluated, based on both effective classification and power consumption. One possible solution to the gunshot detection problem would be to develop one of these algorithms into a sensor mounted with the sensor module. However, this development and production would be far less cost effective

[1]This classification also triggers the generation of an event that is propagated to the event handler agents of the system. See section 3.9 for more information

than the COTS approach that we are going for with this HUD system. Further, the evaluated preprocessing algorithms is compared to a simple absolute value approach, where one simply evaluates the absolute value of the audio signal against a threshold or a signal mean value. This method gives comparable results, in terms of true and false positives, to the more complex algorithms and has a far lower computational cost. The absolute value approach would also be cost effective in a software implementation due to the few computational steps needed to determine if the signal has exceeded the threshold. This would also allow the preprocessing to be implemented in the sensor module itself, thus simplifying the design and workload of the gunshot detection agent, even to the point of making it redundant. In [Freire and Apolinário Jr, 2010] a method based on gunshot event templates is presented that shows good results even in noisy environments while still being computationally inexpensive. Another approach, in [Gerosa et al., 2007], even adds the ability to detect screams, another good indication that a critical situation has risen.

As we can see, there are a wide range of options for detecting many of the critical situations that can arise through ambient audio signal processing. Many of these also requires little processing power, which suits the limited resources of a wearable system well. For our purpose, the simple threshold method should be more than sufficient, though.

3.9 Context based Notifications

[Starner, 2001a] describes the term ”context sensitive reminders” as a key service that could be provided by a wearable system. Here we will dub these context based reminders ”Notifications”. The notifications are textual and graphical events displayed by the GUI in order to inform the user in an efficient and contextual manner. These notifications start out as events triggered by the specialist agents, our classification experts. The notification agent will then generate a notification based on these events, as well as a handler. The handler is basically a choice given to the user on how to act upon the notification. In the simplest form the handler enables the user to dismiss the notification, removing it from the GUI plane and the notifications queue. The notifications queue is the Notification Agents way of prioritizing notifications sent to the user. Each notification is given a priority weight value, within a given interval, a time stamp and a time out counter. New notifications are queued so that higher priority notifications are at the front and, in cases where the priority value is similar, lower time stamps come first. Upon triggering of the time out counter the notification is removed from the queue. This way the notifications will not pile up over time, blocking new incoming notifications. Another measure that will be important in this regard is the ability of the

Notification Agent to limit the generation of duplicate or redundant notifications. This can be achieved through a similarity measure between incoming events and the events triggering the notifications currently in the queue and in recent history. By saving the received events from a given interval before the current time stamp, we can thus limit notifications triggered by similar events to one every n minutes.

Another way of minimizing clutter in the notification queue is to handle minor notifications in a different manner. By utilizing a status area in the GUI we can display these to the user without demanding any actual action on the users part. Examples of such minor notifications might be new message received, Rules of Engagement updates etc.

3.10 Color Coding

In 2.8 we saw through [Ancman, 1991] that perception of colors can be a bit tricky when displayed in the peripheral vision of humans. This has to be taken into account when choosing color coding intended to be viewed peripherally as opposed to in the foveal vision. In other words, we need to use different color coding schemes depending on whether user attention has been attained or not. In the focal view the user can plainly differentiate most colors, and we are therefore free to use any color coding we deem appropriate. As we can see from some of the military terminology in section 1.6, the color blue is regularly used to represent friendly forces. Also, during military exercises, units representing the OPFOR(Opposing Force) is typically marked with red markings or tape. It is therefore natural to use these color codes in our GUI when trying to represent these two categories. That is, blue for friendlies, red for enemies. These color codes will be relevant for displaying map artefacts on the systems navigational aids. Next we need a color coding for neutral map artefacts, such as points of interest etc. There is no military standards for this, that is known to the author, however since a neutral artefact can be used to represent an unclassified entity, yellow seems like a good preliminary choice. The context based notifications described in section 3.9 are all given an importance rating. By using color coding we can allow the user to recognize the importance or unimportance of a displayed notification through a short glimpse. We use Red-Yellow-Green color coding for these notifications, allowing the user to naturally recognize these colors meaning from previous experience.

In the Peripheral vision we have a different set of rules. Here the use of color will be in the grabbing of a users attention. It is therefore important that the attention cue is recognized as exactly that and not some other form of noise. When doing some preliminary testing of the GUI, an attention cue of white screen flashes was used. This was experienced as a set of flimmers by the testee, and not as a definitive cue. Knowing what we know about color recognition in the peripheral vision from section 2.8, it is clear that blue screen flashes would be far more recognizable in the peripheral vision of the user. Another factor is that blue flashes are already color coded for attention grabbing from their use on emergency vehicles for grabbing other drivers attention in traffic.

3.11 Retaining User Control and Building Trust

Even though we want to delegate as many as possible of the menial tasks of the user to the system, we do not wish to take the control of the system away from the user. How should we balance between user control and system autonomy? A key goal for this system is to manage information flow to the user. The information gathering and analysis should be an autonomous process. Actions based on this information on the other hand should at some point go through a user review. The critical nature of information tasks in a high risk environment necessitates the oversight of the user before actuating any form of information propagation. For example, in a given setting, a shot may be fired even though there is no ”Troops in Contact”(TIC) situation relating to that shot. Automatically sending a contact report without user review in this case could result in resources being pulled from areas where they are most needed in order to counter this false threat. In this case one would need to consult the user on whether or not to proceed in sending the contact report. Using explanations the system can give the user an understanding of the reasoning behind the systems actions. This could serve both to help the user understand the capabilities as well as limitations of the system, build his trust in the system as well as giving him the basis for making a decision. This can be done through a simple dialogue with the user. For example, in the event of a gunshot, the following question should be posted by the system through a notification: ”Gunshot detected. Should a TIC-report be generated?” This short message provides the user with two things. First is the rationale behind the generated notification. Next is the options the user has to consider.

Given the fact that the user very well could be in a TIC situation it is reasonable to assume that he is preoccupied with other tasks and has little attention towards what the system is doing. This is where User Attention Management comes in to play in order to attain user attention needed to get the report reviewed and sent.

3.12 User Attention Management

The HUD display is positioned in the periphery vision of the user. We therefore need a way for the system to grab the users attention when required. This could be in the case of an event requiring a user input or decision. Given what we know from 2.8 we know that the peripheral human vision has its limitations when it comes to color vision. In 3.10 we found that a blue screen flash was a suitable cue for grabbing user attention from his peripheral view. Next issue at hand is how to decide when this cue should be used. This task is delegated to the notification agent. By combining the user state and the importance rating of the notifications it generates, the agent can determine when a notification is sufficiently urgent to warrant the attention grabbing cue. Given a heightened emotional state of the user and a critical notification the notification agent might fire several several attention grab cues until the user responds to the notification. Lower importance notifications does not necessarily need an attention grabbing event, allowing the user to review them at his own time.

3.13 Facilitating use of text messaging

Short text messaging has long been prevalent in civilian communications through the widespread use of cellular phones, but this medium of communications has not seen the same rate of adaptation in military communications. This is in part likely due to limitations of the user interfaces employed by military communications technology. As shown by a small survey made in [Moen, 2014] the short messaging service was rarely used, even though the communications technology in question did provide a short messaging service. By facilitating increased use of text messaging we can gain several benefits. First it will relieve the push-to-talk service of unnecessary traffic making it more available for more critical and time sensitive communications. Secondly, text messages can more easily stored and reviewed at a later time, both during the course of an operation, and as part of a debriefing process after the fact. This also means that time is saved by reducing the amount of information having to be written down by hand, which is both a time consuming and impractical process. This is especially true when working outdoors in poor weather. All this means that we need an interface that facilitates the writing and propagating of text messages in an efficient and available manner. Luckily, the bureaucratic nature of most military organizations provides us with some help in solving this problem. In most military settings messages are written in a predetermined form, according to standardized template. A preliminary contact report for instance, might look something like this:

Preliminary Contact Report

Time sent: [When this message was sent]

From: [Who sent it]

To: [The intended message recipient]

When: [Time of contact]

Where: [Senders location at time of contact]

Wounded: [Number of wounded]

This means we have a number of templates that will make generating these texts a trivial problem, given that we have the required information available. This process is similar to the Canned Text technique described in [Lacave and Diez, 2004] where variable values are added to available slots in a predetermined text. In the case of the preliminary contact report we already have most of the information on hand through our sensors and system state. The gunshot algorithm described earlier might provide a time for the first shot fired and the GPS sensor provides the corresponding GPS location. Since the "Time Sent" and "From" fields are trivial we are left with "To" and "Wounded". These can be given a default value, for instance the platoon commander as recipient and 0 as the default number of wounded. The system can then prompt the user to review and modify these values, if he so chooses.

3.14 Navigational Aids

One of the key functionality of this system will be to provide the user with navigational aids that help him orient and make sense of his surroundings. The overall goal being to provide the user with a situational awareness tool based on the extensive flow of data that a pervasive and ubiquitous system can provide. In [Fukatsu et al., 1998] the idea of a birds eye view is presented for navigation in a virtual reality environment. A small part of the users view is dedicated to this navigational aid, providing him with a better understanding of his surroundings. Another example is the situational awareness aid described in [Bell et al., 2002]. Here a 3D representation of the surroundings in an augmented reality setting is displayed on a see through display. The representation rotates according to the users heading so that the orientation of the representation corresponds to the orientation of the user in relation to his surroundings.

In our case the navigational aid is positioned in the peripheral vision, making it less obtrusive to the users field of view. The navigational aid will have two main modes. First is the situational awareness compass that will make out a significant portion of the systems main screen. The compass will have a north marker and rotate according to the users heading, inferred from a compass sensor in the head mounted sensor module. The centre of the compass disk will represent the users location, the coordinates of which will be displayed at some other part of the main display. Elsewhere on the compass disk, map artefacts are displayed showing their distance and direction relative to the users. This allows the user to estimate the distance and direction to these location with a quick glance.

The second mode is the map screen. This screen can be navigated to by the user and will provide a more detailed view of the surroundings. Here the user can zoom in and out, add map markers and waypoints etc, like one would expect from a traditional GPS navigational system. The same map artefacts will be displayed here, including nearby friendly units, points of interest etc.

3.15 The Graphical User Interface

Although the Intelligent User Interfaces comes with their own set of design challenges, we must not forget the traditional GUI design principles when designing the GUI part of the system. The article by Watzman[Watzman, 2003] mentioned in the chapter intro also provides some useful guidelines on usable interface design. The Recon design guide by Recon Instruments provides a good set of tips when designing a GUI for their HUD, [Recon Instruments, 2014]. Specific to the Recon HUD is the need for a padding of unused pixels around the displays edges. The reason why becomes clear when one tries to display information close to the displays edges, as some distortion is experienced in these padding areas. The guidelines also provide suggested fonts and font sizes, and emphasizes the use of the maximum contrast possible when displaying text. In section 3.10 we found that color coding the context based notifications based on their importance let the user review their relative importance through a quick glance. In order to achieve high contrast while color coding the notifications thes colors should be displayed somewhere else than in the text background. Therefore a color coded frame is used around the notification popup elements. Text that provide important information to the user should be displayed with a high degree of contrast throughout the GUI.

Even though text is an important tool for conveying information to the user, visuals are more powerful in that they enable complex information to be displayed in a simple matter. Take the blue force tracking feature and the Situational Awareness Aid for example. When communicating over a voice only medium, two users can only share their own position in absolute coordinates. Both users then have to manually convert the coordinates into a relative position in relation to their own position, before they both have a sense of where the other is located. By using visual tools and the processing power of the wearable computer, these coordinates can be used to create a visual representation of their relative distance and direction from the user. This information can then be absorbed by the user through a quick glance, rather than through the use of a map, compass, measuring stick, time and some brainpower. This is where the color coding comes into play, giving the visuals extra meaning. Yellow popup means medium importance, a blue blip on the Situational Awareness disk means a nearby friend etc. Watzman also supports the claim that visuals are far more effective at explaining data and concepts than text, given that they are of good design. She also claims that this saves screen real estate by taking up less space than text would in explaining the same concepts. The GUI should therefore use visuals as much as possible, if applicable.

Next issue of GUI design is navigation, that is navigation within the GUI. The users direct interaction is typically limited to the 6 button remote control, consisting of 4 directional buttons, a ”select” and a ”back” button. The GUI should be built with this remote in mind, since it will be the users primary input modality. Therefore a GUI map is proposed that is centred on a main view, with the surrounding views being accessible through the directional buttons. Advanced features of each view can be accessed through the ”select” button. This is also in accordance with the Recon Instruments design guide. Here too we can make use of visuals to save space. Since each of the buttons on the remote control is marked with a clear and recognizable icon, these icons can be used when providing the user with options. An example of this would be a notification about an auto generated message that requires user review before being sent. By providing each choice to the user with an icon representing a button on the remote, the user would know immediately through association how to provide a response to the choices.

The GUI map will consist of the following five views:

MainView The standard view of the system. Includes the Situational Awareness Aid, a textual status area, and a notifications status bar.

MapView Pure map screen. Has advanced mapping functions such as adding waypoints, map browsing and other functions you would expect from a GPS navigation system.

MessageView This is where the user can receive, send, read and generate text messages to and from other TacNet nodes.

NotificationsView Here the user can review unread or dismissed notifications in his own time.

ContextualMenuView This is where agents can offer on demand services, such as video capture, face recognition, system settings etc. The idea is that the user knows best when these advanced and power consuming features should be employed. This menu adapts to what services are available on the local node. This menu might in time be able to provide remote services as well.

3.16 Determining User state

In 2.9 we divided context into two sub categories. Internal and external context. That is, external context can be measured fairly easily with our sensors, whereas the internal state of the user cannot be directly measured. The internal user state, therefore must be derived through pre-existing knowledge and external sensor data. In particular we are trying to determine the emotional state of the user, given our findings in 2.7 that emotional arousal can affect the ability of the user to effectively solve tasks and handle information flow. Here we also saw the physiological effects of emotional arousal and how these can be used to classify the users emotional state through a simple heuristic. This means that through a simple heart rate sensor we can partially classify the users emotional state. However, the heart rate by itself can be a very unreliable basis for classification and should therefore be augmented with other modalities. Respiratory rate measurements can help classify emotions better, as we can see through the heuristic learning tree presented in [Rainville et al., 2006]. Apart from using pre-existing heuristics one can infer emotional states from contexts where one could reasonably assume a heightened emotional state. Given a ambient sound sensor and gunfire is detected, one can reasonably assume that the user is in a heightened emotional state, for example.

3.17 Data Security VS Data Robustness

Reasoning and learning processes used by the system necessitates the storage of a great deal of data. In a military setting, this might not be acceptable in the same degree as in a civilian setting. For example, information on the movements and locations of friendly units should be kept away from anyone without a legitimate need for such info at all costs. We therefore need a way to store reasoning knowledge and experience in a way that retains the basis for reasoning but does not provide a blueprint for the users standard operating procedures. An example of how this could be achieved is by deleting some of the raw data after the reasoning and classification relying on that data is completed. Given a GPS data stream, one would typically need only the last couple of GPS readings for most reasoning purposes. Cumulative data such as distance travelled can be iteratively inferred. Such case attributes might be given a lifetime limit of their own, in the cases where the lifetime limit of the KB cases themselves do not provide enough security. The meta data part of the KB also needs to be constructed with this in mind.

3.18 Defining and Determining Context

In order to provide sound advice to the user, the system needs to be able to determine the context, both the immediate context of the user, as well as the wider context, such as the context of other nodes in the tactical net. This relates to the 2nd attribute of ideal wearable computers, described in section 2.2. A wearable computer should detect and model context. Context was described in section 2.9 and devided into two main categories, internal and external context. In our system the detection and modelling of context is a collaborative effort between agents. Even though some agents specialize on external context, while other focuses on the internal, there are no fundamental differences between the agents in regard to context category. The difference lies mainly in how external and internal context is used. External context provide the system with data that is to be presented to the user. The internal context on the other hand allows the system to tweak the presentation of the external context. As mentioned the context determination is a collaborative effort between multiple agents. The specialist agents are responsible for classifying the current context in their individual sphere of expertise. The map agent determines the spatial context, while the emotions agent tries to determine the internal emotional context of the user, etc. This context classification is then saved to the shared KB for other agents to use, and if necessary a contextual event is created, to be handled by the event handler agents.

Chapter 4

System Models

In this chapter a set of system models are defined in order to illustrate how the system is intended to function and be implemented. Special emphasis is put on the agents and their interaction that enables them to solve the systems goals.

4.1 Agents

Following is a list of planned agents and their description. Each of these agents warrants their own separate design process for detailing how algorithms should be implemented, etc.

Network Agent This is a network interfacing agent. It provides encrypted data communication service between system nodes on the TacNet. It has to be implemented for each specific communication device that the system is to be used in conjunction with. This approach is similar to that of the Application Bridge Agent described in [Cohen et al., 1997].

Blackboard Agent Has the responsibility for saving data and updating the systems knowledge base, as described in section 3.6.

Sensor Agent Handles sensor data streams and distributes the data to subscribed agents.

Notification Agent Builds and handles notifications sent to the user as well as handling user responses to those notifications. Also has the responsibility of user attention management.

Shared State Agent Responsible for building shared state data packages and passing them to the network agent for broadcasting.

Map Agent Is responsible for handling all data flow regarding location and map artefacts. As a specialist agent it also generates events and case classifications in the same categories, for use by event handling agents.

Messaging Agent Tasked with facilitating text messaging. Handles incoming messages as well as providing message generation and sending.

Emotional State Agent Classifies the users emotional state based on sensor data and data classifications from other specialists.

Gunshot Agent Implements audio signal processing algorithm primarily to detect any firearm discharge in the users vicinity.

Imaging Agent Handles data from any imaging sensor connected to the system. Can offer services such as image/video recording, face recognition etc.

BLUFOR Tracker Agent Monitors updates from friendly units connected through the TacNet. Generates Blue-on-Blue and TIC events.

TextGeneration Agent Provides text generation services on demand. Implements a set of explanation methods.

ImageProcessing Agent Provides image processing services. Can distribute heavy processing tasks across several TacNet nodes.

GUI Agent Not really an intelligent agent, but more of an agents way to interface with the user interface.

4.2 Agent interaction architecture

Agent interaction is based on subscriptions and on demand services. Each agent starts its life cycle by broadcasting its presence with a list of resources it wants to subscribe to and what resources it can provide. Through confirmations returned from agents who can supply these resources the agent knows what resources are available and what services it can provide in return. As new services become available the providing agents will update their own services provided lists and broadcast any new services that are enabled as a result. We call this process the Setup Phase. The state machine in figure 4.1 shows how this process of service availability and advertisement is facilitated by each agent.

After the Setup Phase has come to a rest the runtime phase starts. In this mode the system will perform its work servicing the user. How and who the agents communicates with depends on what categories they belong to and on the generated

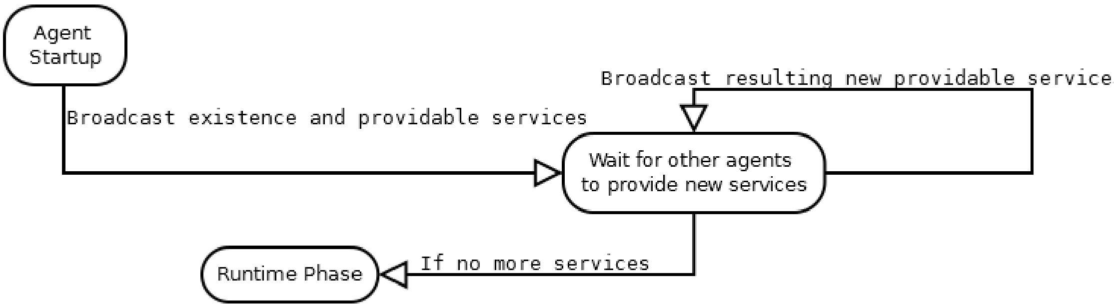

Figure 4.1: Agent startup and yellow pages development procedure

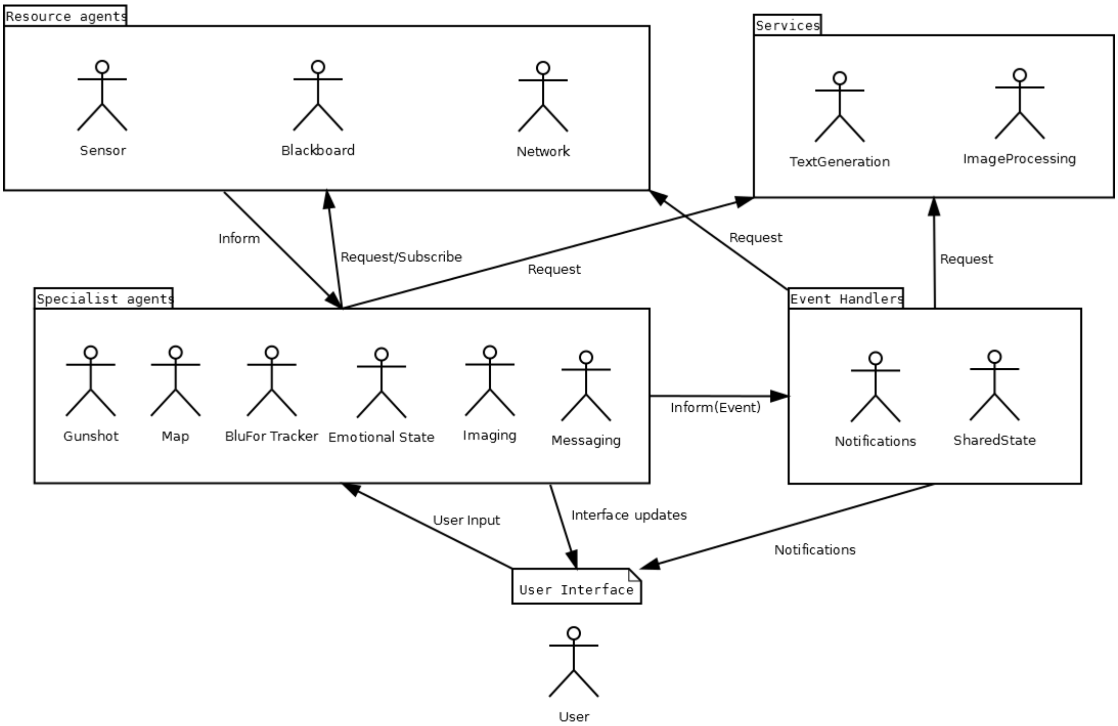

Figure 4.2: Agent runtime interaction, by category

lists. Figure 4.2 shows how members of different categories communicates and works with each other in order to provide the IUI experience to the user.

4.3 System Iterations

In [Moen, 2014] the idea of several system iterations was described. This model had two goals. First was to make developers able to develop the system incrementally, ending up with a working system after each increment. This would enable testing of a fully functional system at an early stage of development. The next goal was to illustrate the adaptability of the system as a whole, where each iteration represents a system setup suited for a different situation. For example, the first iteration would provide just basic navigational aids, whereas later iterations would provide more advanced features such as physiological monitoring and blue force tracking. Figure 4.3 shows how the first iteration of the above interaction architecture might look like.

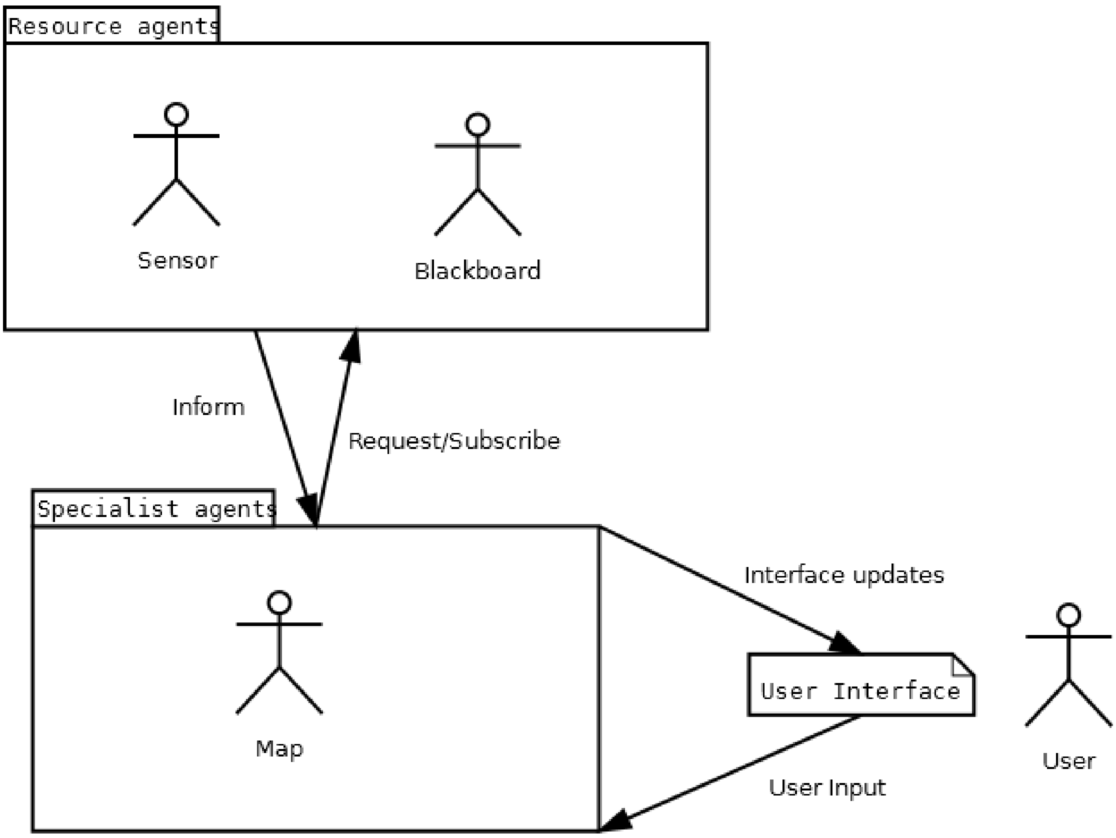

Figure 4.3: First iteration version of the agent interaction architecture

Further description of this iterative approach can be read in [Moen, 2014].

4.4 Scenarios and use cases

This section describes a set of scenarios focused on some key functionality envisioned for the system. In [Moen, 2014] a more anecdotal set of scenarios describing the more overall vision for this system can be read. It follows a military unit conducting a realistic operation against a realistic object, and shows how the HUD system can aid the soldiers in performing their tasks, from the planning stage, through execution and up to debrief and after action report. In the following scenarios, the functionality and reasoning process of the system will be in focus. First a short textual description of each scenario is presented before they are visualized in a use case diagram.

4.4.1 Scenario 1

A military patrol makes contact with the enemy and gunfire ensues. The system generates a preliminary contact report for review by the user. The user looks over the report and approves it for transmission. The system then sends the message to the Tactical Net nodes of the rest of the unit, including the unit commander. The scenario is illustrated by Use Case in figure 4.4.

4.4.2 Scenario 2

Upon receiving a contact report from the TacNet node of another unit member the system brings this to the users attention through the blue flashes cue and a notification. In addition a map artefact is added to the navigational aid denoting the location where the contact occurred. The scenario is illustrated by Use Case in figure 4.5.

4.4.3 Scenario 3

User finds himself in a stressful situation and suffers the effects from his intense emotional state. The sensors of the system detects the physiological effects from this state and the Emotions Agent classifies the user state accordingly. As a result the system information flow to the user is minimized to only the most critical updates, and blue flashing cues from the peripheral display are used to bring these situational updates to the users attention. The scenario is illustrated by Use Case in figure 4.6.

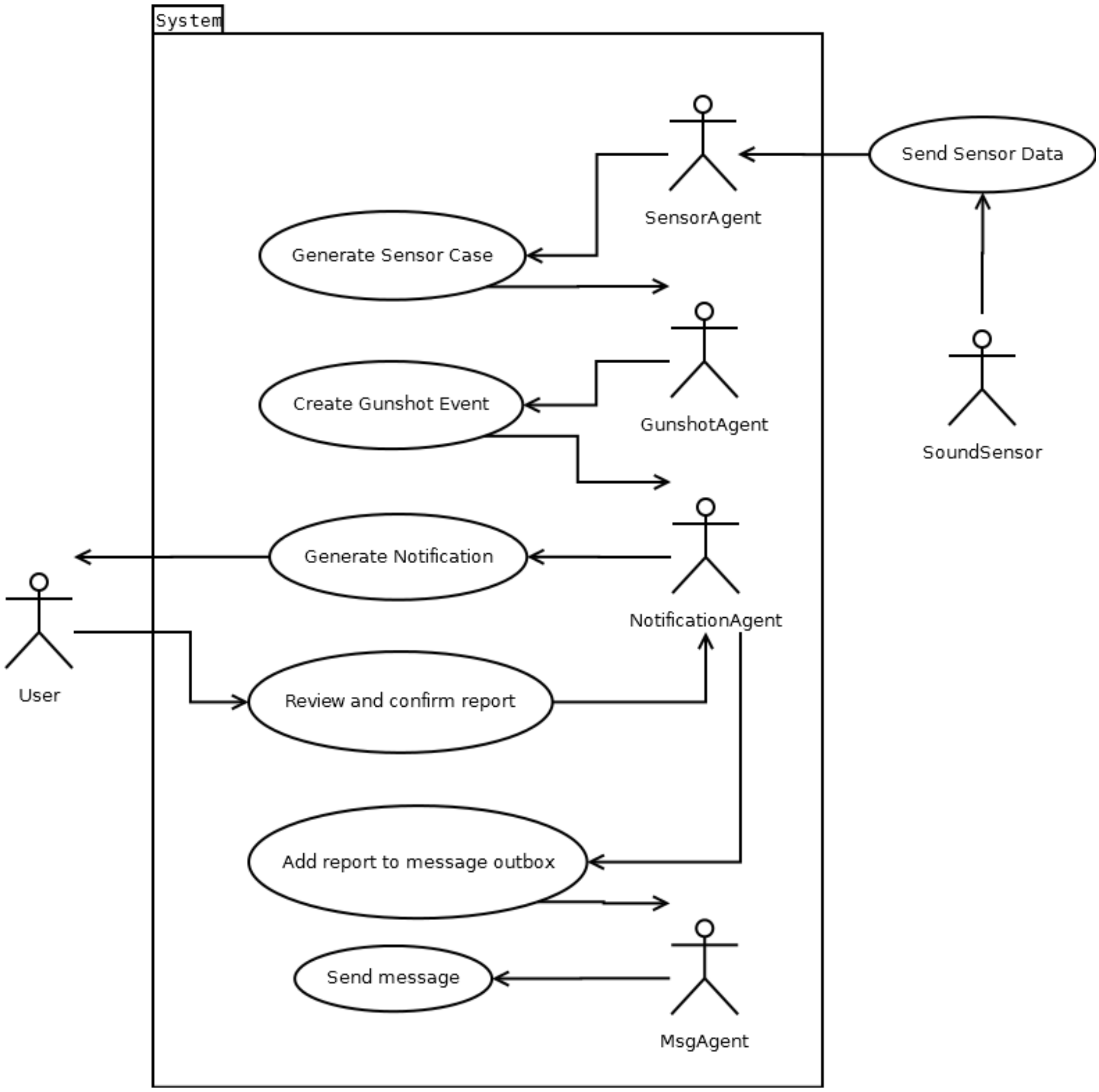

Figure 4.4: Use Case illustrating scenario 1

4.4.4 Scenario 4

While several soldiers are moving through a complex environment each soldiers display is continuously updated with the location of, or direction and distance to, all friendly units nearby set up with a TacNet node. This is displayed visually so the soldier can review this information quickly and gain an improved situational awareness as a result. The scenario is illustrated by Use Case in figure 4.7.

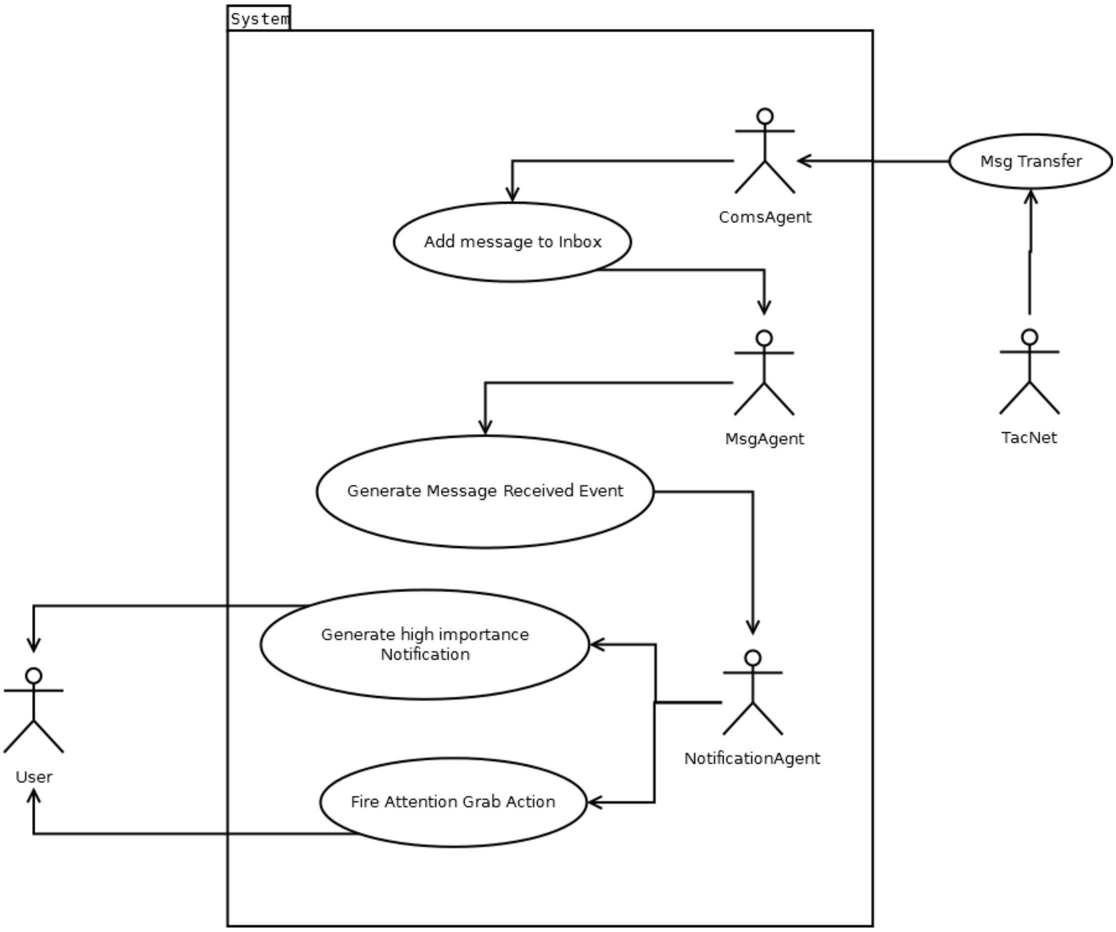

Figure 4.5: Use Case illustrating scenario 2

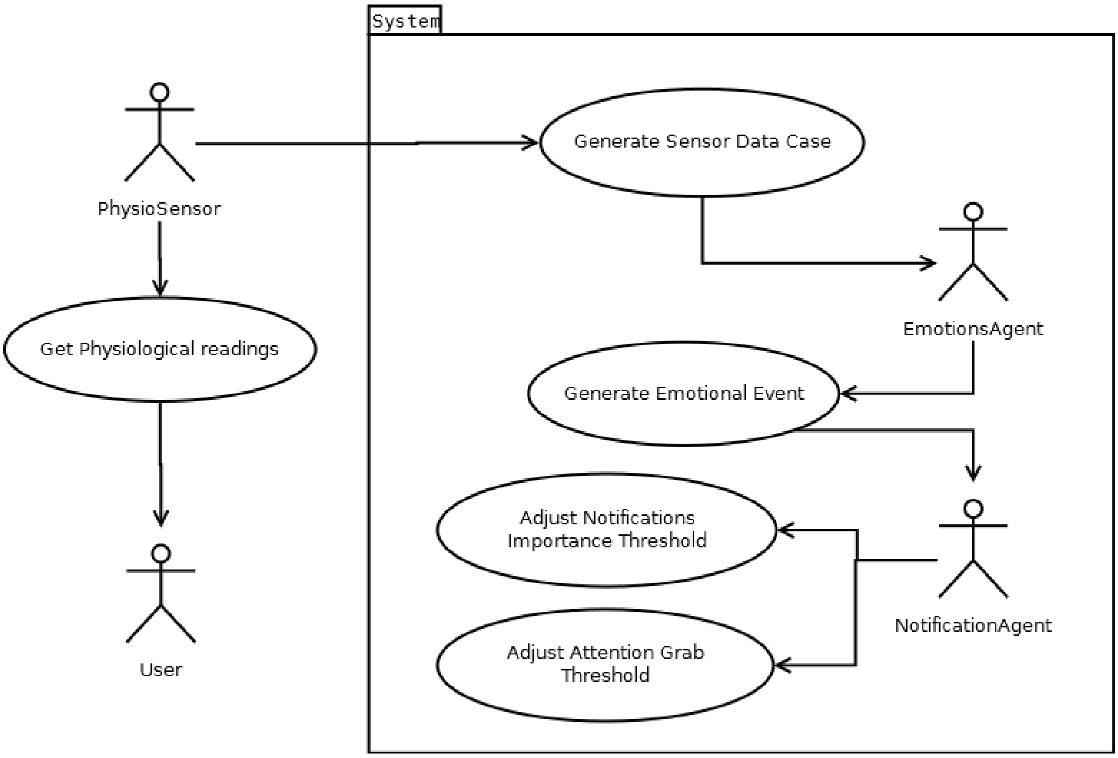

Figure 4.6: Use Case illustrating scenario 3

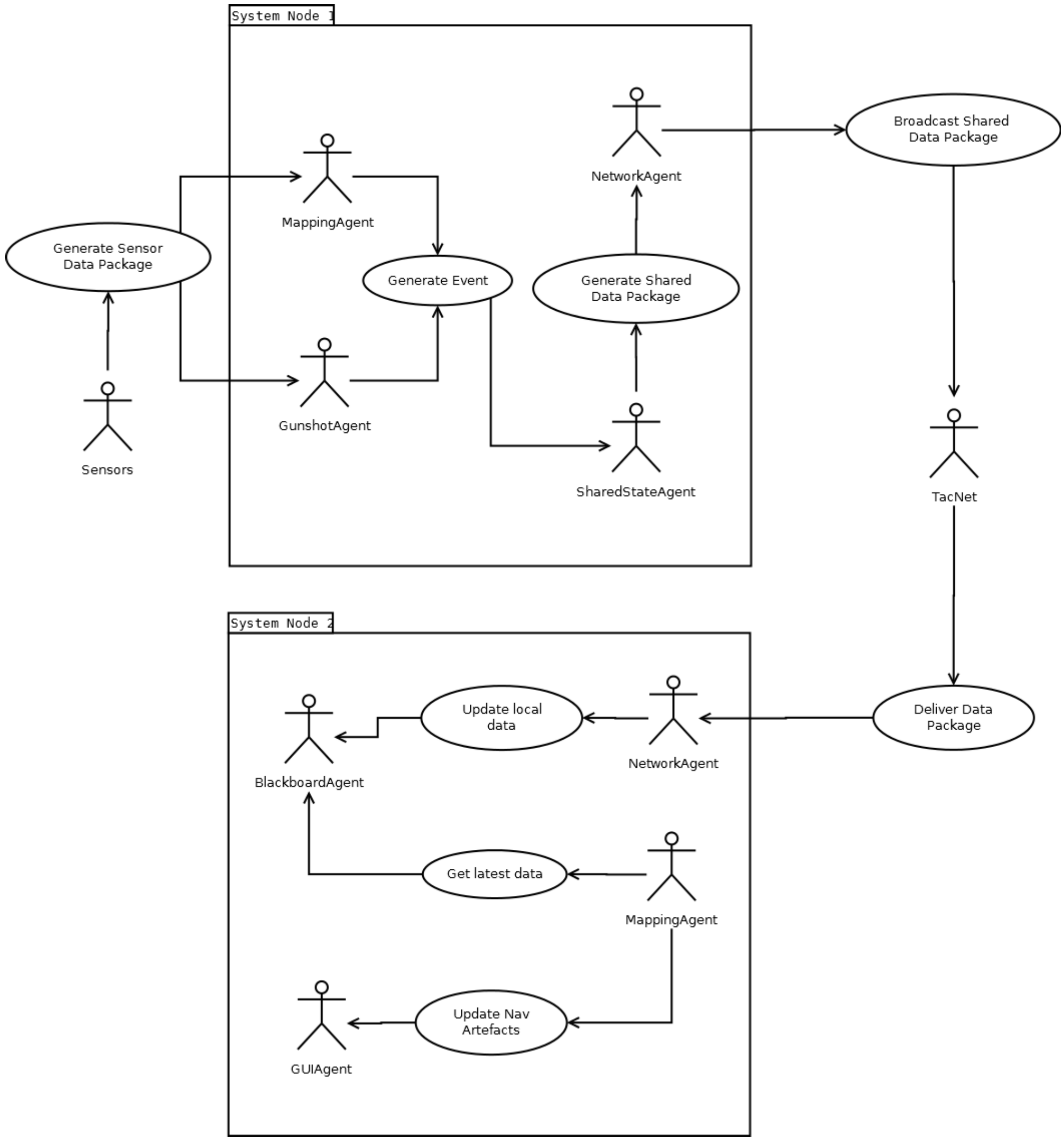

Figure 4.7: Use Case illustrating scenario 4

4.5 GUI Design

The GUI has 5 main views, with the main view being the root view. The 4 other views can be accessed from the main view through the directional buttons on the bluetooth remote control. Being the root view, the main view can always be reached by pushing the back button a number of times or pressing it for a few seconds. Figure 4.8 shows the GUI navigation model.

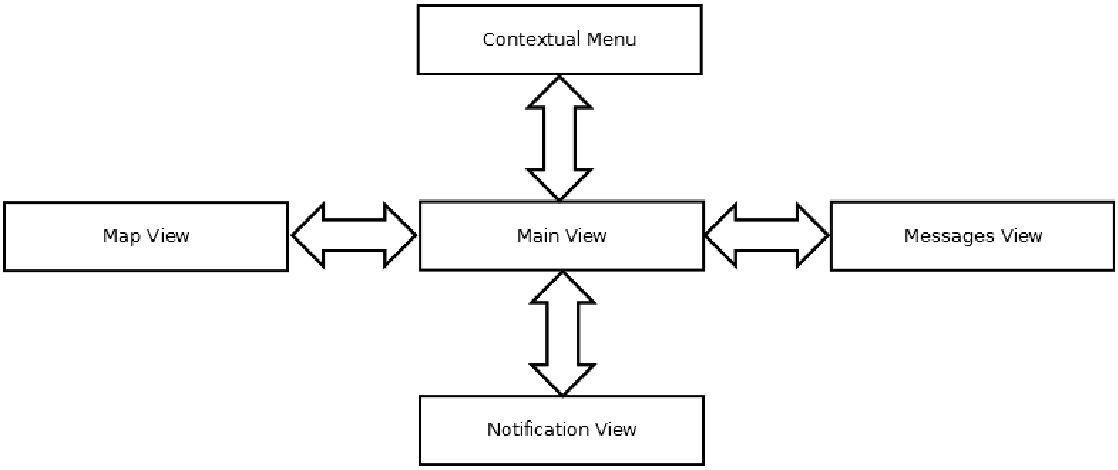

Figure 4.8: The GUI views navigation model

The main view, figure 4.9, consists of three sections. The situational awareness aid or the compass disk, the status area and the notification bar. The red dot on the Situational Awareness Aid indicates north.

Next is the MapView, figure 4.10, and MessageView, figure 4.11. These are special purpose views controlled by the Mapping and Message Specialist Agents. They are only displayed on the users command. The notifications will pop up on top of whatever view is active at the time. Here shown displayed on top of the MainView, in figure 4.12. Notice the use of a high contrast text backgroud, while a colored frame provides the importance color coding.

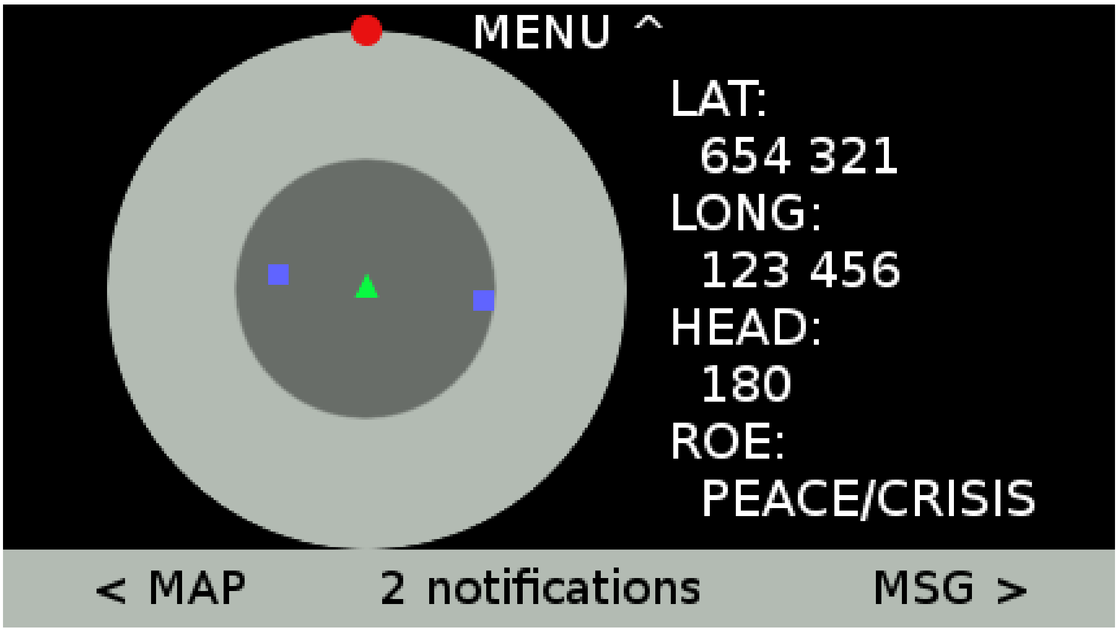

Figure 4.9: The MainView

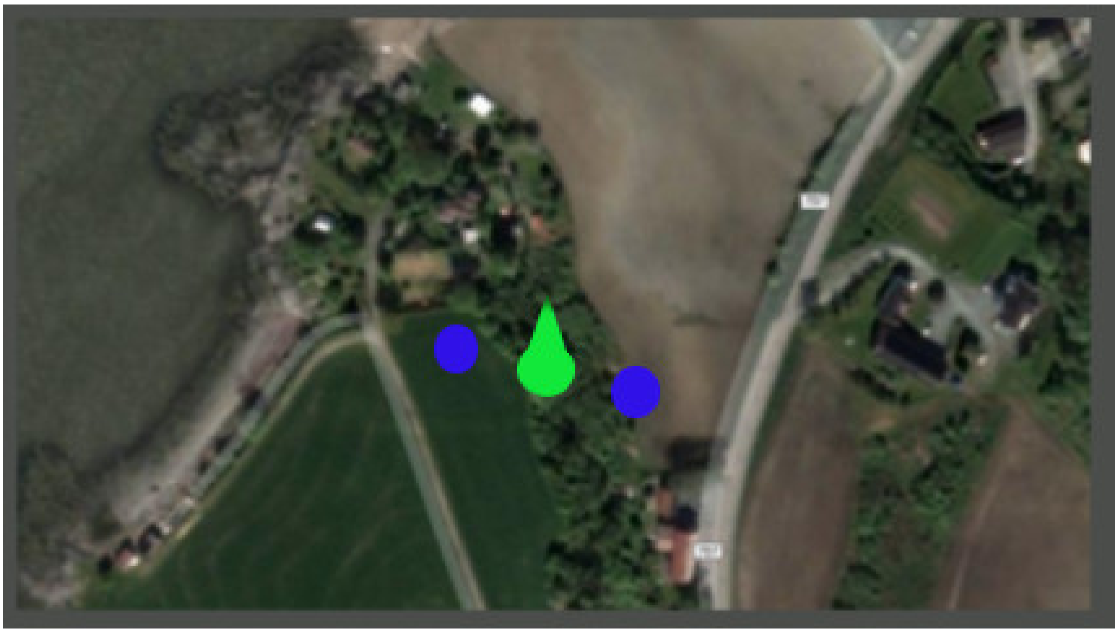

Figure 4.10: The MapView

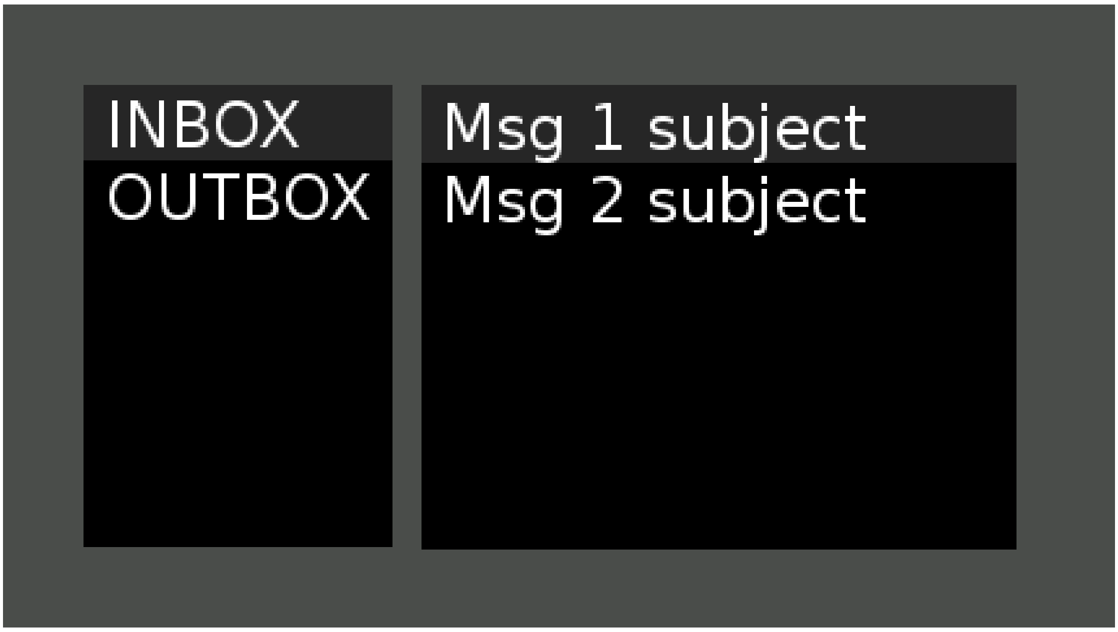

Figure 4.11: The MessageView

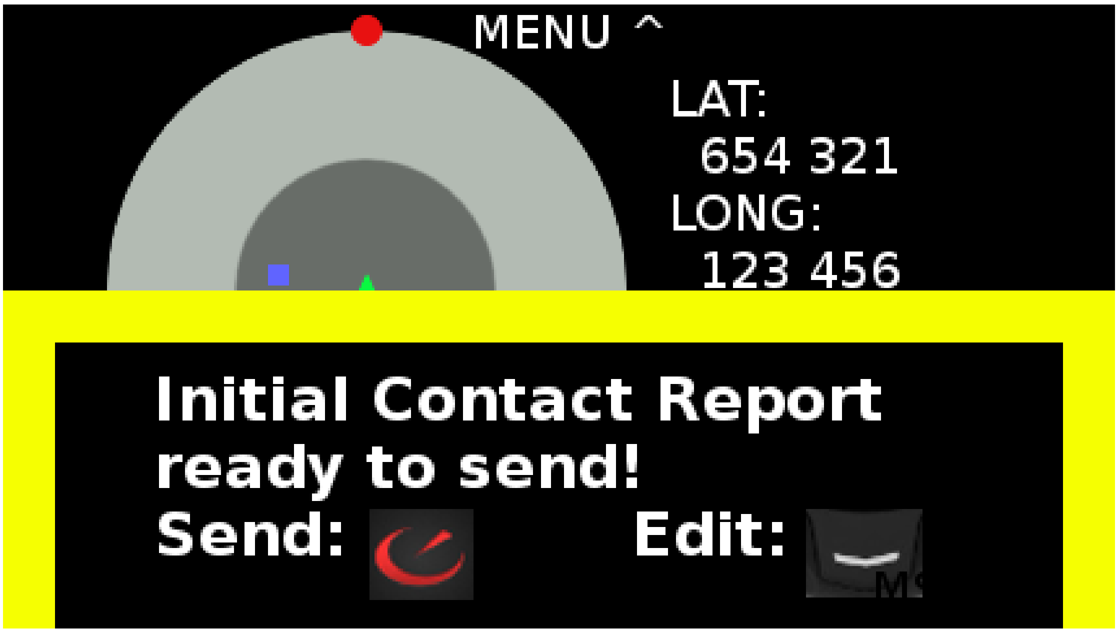

Figure 4.12: A Notification is displayed over the MainView

Chapter 5

Evaluation and Conclusion

5.1 Evaluation

This study turned out to be very broad. As a result there was not room for as much in-depth studies on each issue as one would have wished. The classification of the users emotional state for instance could warrant a whole study of its own. The thesis has instead taken a broad look in order to cover the whole of the system vision upon which it is based.

Other issues is that in many areas in depth studies could not be found. Especially studies that shows limitations of new technologies are difficult to find. An example of this is studies of Augmented Reality interfaces and their ability to meet users expectations. Specifically, in this context, what accuracy a user will expect from AR annotations/objects versus artefacts on a compass disc on a LCD screen.

Lastly, the complexity and broadness of the concept presentation has limited thorough user testing of the concepts. This means that the feasibility of the system is based largely on a theoretical basis and not actual user feedback.

The thesis does however paint a relatively complete picture of the vision and the solutions that can solve many of the challenges that such a system will be faced with.

5.2 Summary

From the theoretical basis presented in chapter 2 we have shown that an agent based approach to intelligent user interfaces is the way to go when building a wearable, pervasive and ubiquitous system. Additionally we get the modularity that the agents provide, enabling iterative and cost effective development and maintenance. As a result the system can be tailor made to a variety of use environments with little effort. This also simplifies the implementation of complex features. The building of support for complex calculations can be done independently from the development of the system as a whole

In chapter 3, five criteria of good design was presented, and design decisions was made in order to meet these criteria. Three of these criteria, appropriateness, durability and cost effectiveness, it has been shown, has been largely met through the modularity and adaptability of both the hardware and software components of the system. Modularity are appropriate for a military setting given the wide range of use environments and thus the need for adaptability. Through the iterative approach we have ensured the durability and cost effectiveness by enabling low development and maintenance costs. The last two criteria, verifiability and impact has not been thoroughly dealt with, since no user testing has been carried out in this study.

The feasibility of specific features of the system has been shown, such as classification of user emotional state, ambient audio events and attention management from the users peripheral view. The thesis has argued that traditional LCD HUDs are still relevant even though the new paradigm of augmented reality is gaining ground in the industry.

Lastly a set of use cases and an agent architecture has shown how the system will function at runtime. Also, a set of interface sketches has been presented in order to show how the system will look like from the users point of view.

5.3 Discussion

So, looking back to the research questions posted in section 1.3, Have these questions been answered? Lets go through them one by one, starting with the human factors.

What human factors needs to be considered when interfacing with a human user operating in high risk situations? We have looked at two main human factors, vision and emotions. We have found that human vision has its limitations, such as color recognition in the peripheral view. It has also been shown that this and other factors are affected by the users emotional state. The inverted-U function has been used to visualize this effect. That is, a decrease in performance can be seen, generally, at very low or very high states of emotional arousal. Even though it has been argued that this function is an overly simplistic view on the effects of emotional arousal on human performance, it serves as an adequate visualization tool for these effects. In addition to vision, information processing and other mental abilities are also similarly affected.

How can a pervasive system in a high risk setting recognize and classify context? Specifically, detect critical situations and detect and consider implications of user emotions. We have shown that agents can work cooperatively in order to classify the current context from sensor data. Specifically identifying critical situations that might affect the users ability to handle information. It has been shown that gunshots and user emotional state can be classified through known methods implemented in specialist agents that are each delegated one of these specialist tasks. Further contextual information can be generated by adding more agents and hardware modules to the system. The information built by these specialists can then be used in the reasoning processes of other agents, which leads us to the next research question.

How can context be used to inform reasoning processes? Here too, the specialist agents are key. Through each specialist agents perspective we get a different classification of the data. Each classification can in turn be used in different reasoning processes. For example, a map data classification will be used to generate content, while an emotions classification will be used primarily to determine how and if that content should be displayed. In addition we have the specialist generated events, that are consumed by the event handlers. This represents the reactionary reasoning processes, reacting to current events, as classified by the specialist agents. The shared knowledge base, known here as the Blackboard Agent, also plays an important part in maintaining a shared world view among agents, enabling an enhanced level of cooperation in reasoning between them. Next issue is then how to use the results from these reasoning processes to the benefit of the user.

How can these reasoning processes and their results be used to create an increased situational awareness within the user? Three main sources of situational information is presented to the user. First is the Situational Awareness Aid on the main view, providing the user with spatial situational awareness. Next is the event triggered contextual notifications. These updates the user on current contextual events detected by the system. Last is the browsable views, including the map and message views. These provide the user with on demand information and services in specific categories.

Together, these three information channels should provide the user with a situation appropriate flow of information. First by providing a continually updated image of the current context. Second by providing situational updates on detected events, as well as offering a choice in suitable responses to these. Lastly, the user can look up the specific information or service he needs by browsing the user interface.

We see that through agent based reasoning we can provide the user with an adaptable intelligent user interface system that provides pervasive contextual services and information in a natural and effective way. Even though the system is planned as an COTS hardware based system, some fabrication of components will be needed, specifically the sensor module and hardware interfaces to the communications technologies it is supposed to utilize for data communications. The sensor module might become more and more redundant as the COTS hardware implements more internal sensors. A new version of the Recon HUD has already added a camera to its internal sensors, and more are certain to be added later. The COTS approach will therefore only become more economical in time. Given the conclusions in section 2.5 on the limitations of the state of the art augmentet reality technologies, further savings can be made by choosing more affordable and tested hardware components. This design is meant as a proposed system to be implemented and used. Some choices has therefore been made out of practical reasons, more than state of the art research. An example of this is the use of Heads Up Displays instead of the state of the art Augmented reality and a simple threshold algorithm used for gunshot detection. These choices was made due to practical implications of the user environment and the limited resources available.

The most key issues not yet dealt with relating to the system design are the meeting of the last two criteria from chapter 3, verifiability and impact, since no thorough user testing has been performed at this time. This will be important in future work.

5.4 Impact

As this is a feasibility study and vision clarification, it does not impact the state of current research in any meaningful way. It does however clarify the vision of the system presented, and thus will work as a good starting point for implementing it, given extensive user testing of the design. A few stands has been made that may seem to go in a different direction than many other studies, such as sticking with the older Heads Up Display paradigm in the face of new AR technology and using old and simple methods for gunshot detection and text generation.

5.5 Future Work

As mentioned, the design needs to be evaluated through user testing, which should be the first next steps. Upon a confirmation and possible refinement of the design, a first iteration system should be built. This should then be subjected to further testing. Continual iterations of development and testing should then be performed, gradually adding to the set of features, agents and hardware components. Implementation of each new component or agent will require a design process of its own.

Bibliography

Aarts, E. and Wichert, R. (2009). Ambient intelligence. In *Technology Guide*, pages 244–249. Springer.

Ali, S. I., Jain, S., Lal, B., and Sharma, N. (2012). A framework for modeling and designing of intelligent and adaptive interfaces for human computer interaction. *International Journal of Applied Information Systems (IJAIS) Volume*.

Ancman, E. G. (1991). Perceptual limitations of peripherally displayed colors on crts. Technical report, DTIC Document.

Arai, F., Fukuda, T., Yamamoto, Y., Naito, T., and Matsui, T. (1993). Interactive adaptive interface using recursive fuzzy reasoning. In *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*, pages 104–110. IEEE.

Bayardo Jr, R. J., Bohrer, W., Brice, R., Cichocki, A., Fowler, J., Helal, A., Kashyap, V., Ksiezyk, T., Martin, G., Nodine, M., et al. (1997). Infosleuth: agent-based semantic integration of information in open and dynamic environments. *ACM SIGMOD Record*, 26(2):195–206.

Bell, B., Höllerer, T., and Feiner, S. (2002). An annotated situation-awareness aid for augmented reality. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 213–216. ACM.

Bittel, J. (2013). On today's battlefield, is information more important than firepower.

Caire, G. (2012). Jade tutorial jade programming for android. *Independently released*, 1(1):1–10.

Chacon-Rodriguez, A., Julian, P., Castro, L., Alvarado, P., and Hernández, N. (2011). Evaluation of gunshot detection algorithms. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 58(2):363–373.

Christianson, S.-Ã. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, 112:284–309.

Cohen, P. R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., and Clow, J. (1997). Quickset: Multimodal interaction for distributed applications. In *Proceedings of the fifth ACM international conference on Multimedia*, pages 31–40. ACM.

Ekbia, H. R. and Maguitman, A. G. (2001). Context and relevance: A pragmatic approach. In *Modeling and Using Context*, pages 156–169. Springer.

Freire, I. L. and Apolinário Jr, J. A. (2010). Gunshot detection in noisy environments. In *Proceeding of the 7th International Telecommunications Symposium, Manaus, Brazil*.

Fukatsu, S., Kitamura, Y., Masaki, T., and Kishino, F. (1998). Intuitive control of ”bird’s eye” overview images for navigation in an enormous virtual environment. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 67–76. ACM.

Gerosa, L., Valenzise, G., Tagliasacchi, M., Antonacci, F., and Sarti, A. (2007). Scream and gunshot detection in noisy environments. In *15th European Signal Processing Conference (EUSIPCO-07), Sep. 3-7, Poznan, Poland*.

Hecht, E. (2002). *Optics*. Addison-Wesley, San Francisco, CA.

Heinroth, T., Kameas, A., Pruvost, G., Seremeti, L., Bellik, Y., and Minker, W. (2011). Human-computer interaction in next generation ambient intelligent environments. *Intelligent Decision Technologies*, 5(1):31–46.

Ishizaki, S. (1996). Multiagent model of dynamic design: visualization as an emergent behavior of active design agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 347–354. ACM.

Kaufmann, H. and Dünser, A. (2007). Summary of usability evaluations of an educational augmented reality application. In *Virtual Reality*, pages 660–669. Springer.

Keinan, G. (1987). Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology*, 52:284–309.

Kofod-Petersen, A. (2007). A case-based approach to realising ambient intelligence among agents. *Department of Computer and Information Science*, page 322.

Kortuem, G., Bauer, M., and Segall, Z. (1999). Netman: The design of a collaborative wearable computer system. *Mobile Networks and Applications*, 4:49–58.

Lacave, C. and Diez, F. J. (2004). A review of explanation methods for heuristic expert systems. *The Knowledge Engineering Review*, 19(02):133–146.

Macias, A. (2013). Supplier denies that the new us military helmet is based on 'halo'.

Maes, P. et al. (1994). Agents that reduce work and information overload. *Communications of the ACM*, 37(7):30–40.

Maybury, M. (1998). Intelligent user interfaces: an introduction. In *Proceedings of the 4th international conference on Intelligent user interfaces*, pages 3–4. ACM.

Moen, G. (2014). Soldierhud. Fall project upon which Gjøran Moens Master Thesis is built.

Pugliese, D. (2013). Firms battle to build future warrior.

Rainville, P., Bechara, A., Naqvi, N., and Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, 61(1):5–18.

Recon Instruments (2014). Design guidelines.

RheinmetallAG (2013). Rheinmetall delivers gladius soldier system.

Roorda, A. and Williams, D. R. (1999). The arrangement of the three cone classes in the living human eye. *Nature*, 397:520–522.

Singh, S., Puradkar, S., and Lee, Y. (2006). Ubiquitous computing: connecting pervasive computing through semantic web. *Information Systems and E-Business Management*, 4(4):421–439.

Sistemika, M. (2014). Personal eye system.

Sørmo, F., Cassens, J., and Aamodt, A. (2005). Explanation in case-based reasoning–perspectives and goals. *Artificial Intelligence Review*, 24(2):109–143.

Starner, T. (2001a). The challenges of wearable computing: Part 1. *IEEE Micro*, 21(4):44–52.

Starner, T. (2001b). The challenges of wearable computing: Part 2. *IEEE Micro*, 21(4):54–67.

Urke, E. H. (2014). See the norwegian armed forces driving with oculus rift.

Watzman, S. (2003). Visual design principles for usable interfaces. *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*, pages 263–285.

Weiser, M. (1991). The computer for the 21st century. *Scientific american*, 265(3):94–104.