Chapter 6 – ADVANCED UMV OPERATOR INTERFACES

Chapter Lead: M. Draper

Contributors: T. Barry, G. Calhoun, J. Clark, M. Draper, M. Goodrich, C. Jansen, J. Kessens, F. Kooi, A. Lefebvre, S. Murray, J. Nelson, C. Nielsen, G. Osga, A. Oudenhuijzen, M. Quigley, R. Shively, B. Simpson, R. Stone, L. van Breda, J. van Delft, J. van Erp

6.1 INTRODUCTION

Even with rapid advances in computer processing, automation technology, and artificial intelligence methods, there remains a critical need for human involvement in order for UMVs to successfully perform their missions. The human provides unique strategic and innovative decision-making capabilities within complex, dynamic, and time sensitive situations. UMV operator performance and, by extension, the UMV operator control/display interface, will be even more critical to achieving anticipated new and increasingly complex UMV capabilities including close-coupled operations with manned systems, UMV interoperability, and military strike/combat operations.

Given that humans are to remain a key component of UMV systems for the foreseeable future, it is important to recognize the unique challenges levied upon the operator. These challenges include the effects of system time delays (both fixed and variable), bandwidth limitations (which can be intermittent), datalink degradations/dropouts, and the loss of the rich supply of multi-sensory information often afforded to onboard operators. With future highly automated UMV systems, issues also include functional allocation of tasks between the operator and the system, human vigilance decrements, 'clumsy automation', limited system flexibility, mode awareness, trust/acceptance issues, failure detection, and automation biases. Additional challenges have been documented in detail elsewhere (Chapter 2). However, it is also important to note that the physical separation of crew from vehicle might also offer some unique benefits that should be exploited. Besides the obvious benefit to crew safety, it is quite likely that available bandwidth and the variety of available information sources might be, in certain cases, far greater for a geographically-separated UMV crew versus an onboard operator, potentially resulting in more situation awareness rather then less. This, of course, assumes that a well-designed operator interface exists that can rapidly filter and fuse this expanded information into intuitive displays, again underscoring the need to attend to operator interface issues to ensure maximal system performance.

It is also important to note that as technology advances, the role of the UMV operator must change as well. Therefore, UMV operator interfaces should not be considered 'one-size-fits-all', but must be tailored to match the capabilities and limitations of the host system and intended mission. Most current UMVs require that operators have the capability to manually control the vehicle and activate state changes (i.e., direct tele-operation). Thus, operator interfaces for these vehicles can best leverage the numerous lessons learned from decades of inner-loop control design research, while applying novel interfaces to combat challenges that are uniquely associated with UMV operation.

With new, highly automated UMVs, the operator's role is becoming more supervisory in nature, overseeing the automated activation of programmed events (e.g., making sure the appropriate event is activated at the appropriate time), managing changes to the automated mission plan, and making more strategic-level decisions. These operator interfaces must take into account issues associated with automation management, including vigilance effects, brittle/clumsy automation, sudden workload spikes, etc.

Continuing this trend beyond the current state-of-the-art, a vision exists for a new interface paradigm for controlling next generation UMVs. This envisioned interface system involves multiple semi-autonomous UMVs being controlled by a single supervisor. These UMVs will have the capability to make certain higher-order decisions, independent of operator input and pre-defined mission plans. This capability of the UMV 'to decide' constitutes a whole new set of challenges for operators, as they will be required to rapidly judge the appropriateness of these decisions and assess their impact on overall mission objectives, priorities, etc. Future operator interfaces will need controls and displays tailored for multi-UMV control and must allow the operator the capability to easily inspect/override the autonomous UMV decision-making logic. These interfaces will also need to provide information fusing/filtering algorithms, intelligent prioritization/cueing logic, and possibly some form of adaptive task allocation in response to rapidly changing events and/or workload levels.

This chapter explores many relevant issues surrounding UMV control/display interface technology. The chapter begins with a discussion of the importance of pursuing a user-centered design methodology in designing UMV operator interfaces. This is followed by a section detailing available and applicable design guidelines as well as noticeable gaps in this area. Next, several candidate data input and display technologies are discussed in turn. Each technology section (with few exceptions) begins with a summary description of the technology area along with any commonly accepted subcategories of that technology. Next, there is a discussion of actual or potential application of the technology to UMVs. Lastly, technology maturity, challenges, and unresolved issues are summarized. The chapter then focuses on operator interface issues associated with multi-UMV supervisory control and concludes with additional UMV interface considerations such as scale, mobility, and various communication constraints.

6.2 USER-CENTERED DESIGN FOR UMV CONTROL

UMVs operating in collaborative, semi-autonomous groups represent a major advance in military capability. Multiple remote platforms conducting operations in hazardous environments – where errors can have serious military and political consequences – also represent a new level of engineering complexity. The missions planned for UMVs are diverse and complex by any standard; the control logic required to operate them will also be complex and will always contain an element of risk.

Because UMVs, like all military systems, are extensions of human intent, UMV effectiveness can only be as good as the clarity of that intent. That intent is conveyed via the primary tool for remote human control – the user interface. Good control depends on good information – which must be complete, sensible, reliable, and timely. The quality of the user interface, therefore, will ultimately determine the quality of system performance. A good interface can help the human controller to maintain situation awareness of the remote environment where UMVs are operating, can enhance goal setting under dynamic vehicle and mission conditions, and can optimize human analytical processes and decision-making.

Certainly, a lot is known about how to provide the human controller with good information and decision support tools – the challenge is to apply this knowledge aggressively and early in the design process. Human factors engineering (HFE) methods inform the interface design, i.e., its content, layout, and interaction metaphor, while human-system integration (HSI) methods place the interface in the context of the total human-UMV system including its missions, operating environment, and support requirements.

Human factors engineering defines the form and behavior of the controller interface. It applies knowledge of human perceptual capabilities, motor skills, memory capacity, decision-making processes and communication

styles to the design of interface functions. Complex applications involving teams of operators must also address human characteristics related to shared human goals, team interactions, and distributed decision-making. The primary objectives of HFE is to ensure that information presented to a controller matches their cognitive and decision-making requirements in real time, that the information is easily understood, and that controller decisions are easily implemented.

Human-system integration is a critical component of the systems engineering discipline that addresses human capabilities and limitations throughout the UMV development process. The objective of HSI is to match technology to user capabilities in order to optimize mission performance. HSI also employs HFE methods and data, but applies them to broader issues of safety, training, and to technology or mission evolution over the system life cycle. "Integration" implies a task conducted amid multiple, often-conflicting system design demands so HSI must address such issues as the impact of network communications architectures, varying levels of UMV automation capability (including intra-vehicle communications and decision-making), and the human role in coping with degraded conditions. HSI methods can ensure that systems are built to accommodate the characteristics of the personnel who will operate, maintain, and support them.

HFE and HSI tools and methods are highly complementary and, insofar as they reflect advocacy of the human role in system design, may be combined into an overarching concept of user-centered design (UCD). The UCD approach reflects an appreciation of human potential in the system engineering process. Regardless of the engineering sophistication of any military system, human controllers – with their unique abilities to sense and understand both the system and the mission – have always added the critical margin of performance that makes such systems viable. The UCD orientation to UMV development, therefore, represents a critical contribution to the success of such complex systems.

The new capabilities reflected by UMV systems will stress current engineering design practice regarding control logic, communications networks, and multi-level automation. These same capabilities will stress current perspectives of how human presence can best support the new missions and expanded mission modes enabled by UMVs. UCD, therefore, will be faced with new challenges that will extend our concepts of human-system interaction to settings where machines are cast (at a behavioral level) more as colleagues/ collaborators than as mere tools in military operations:

- New models of human-machine and human-human communication must be elaborated, especially in the context of network-centric warfare doctrine.
- New approaches to smoothly and safely altering multi-agent control configurations (the consequence of enhanced automation capabilities) must be defined.
- Because UMV control should be achievable anywhere – from the fixed-base command center to the mobile, individual warfighter – scaleable interface designs, based on formal HFE/HSI principles, must be generated.
- As hardware and mission concepts change over time, UCD methods must accommodate the certain requirement for graceful (rather than disruptive) life cycle evolution.
- Finally, new testing and validation methods must be generated and proven to ensure that the resulting systems will perform as required.

UCD has a clear and historically proven role in good system design. Furthermore, when explicitly included at the beginning of a sound systems engineering program, UCD can be accommodated without interference to other engineering tasks. While UMV system design must address new challenges across a broad front of engineering specialties, UCD integration is an enhancing – rather than distracting – step in such an ambitious

undertaking. An irony of many military acquisition programs is that UCD is never deleted; some consideration to HFE and HSI issues is almost invariably required either at the time of system fielding or shortly thereafter, when performance problems force an examination of the human role in mission effectiveness – usually with increased cost or delay. UMV design can expand both our engineering and human interface design knowledge, and can add to the technical "toolboxes" of both domains. To realize the full potential of the UMV concept, however, UCD design methods must be fundamental to the engineering process from its initiation.

6.3 GUIDELINES AND GAPS

Since the advent of highly capable uninhabited vehicles, notably in the application domains of defence, offshore oil and gas exploration, attention has increasingly focused on the development of technologies necessary to endow vehicles with complete autonomy. This approach has not met with widespread success. Evidence frequently points to the fact that the human operator still has a significant role to play in the future of uninhabited vehicles, as part of a control continuum that ranges from direct tele-operation during critical mission phases and recovery modes to the high-level supervision of single or multiple platforms. However, few (if any) usable guidelines and/or experimental test beds exist to help ensure that human factors issues are taken into account early in the design lifecycle of uninhabited vehicles and their control-display interfaces with the human operator or supervisor.

6.3.1 Pre-Requisites to Selecting Human Interface Devices

The market for off-the-shelf interface devices worthy of consideration for UMV application is rapidly growing. In particular, the exploitation of games console hand controller designs is gaining pace, as there is a perception (unproven at the time of writing) that new recruits to the Armed Forces will adapt "seamlessly" to advanced military interfaces if those interfaces are based on familiar gaming devices. Despite this, an early human-centred design approach to device specification/design/procurement is essential; any display or control device must be selected on the basis of a thorough review of the tasks expected of the human operator, regardless of whether his/her role is that of flight/mission/payload specialist, supervisor or field dispatch (see Section 6.2).

The pitfalls of being swayed by technology push are now well documented throughout the international human factors community. Obsolete and unused Virtual Environment Centres bear witness to the hazards of selecting technologies before the needs of the end user population have been fully understood. Yet still one finds examples where, whilst the information to be displayed to the human user has been subjected to strong human factors principles, those same principles are ignored in what is blatantly a prescriptive selection of unproven human interface technologies. This situation is particularly rife in the virtual reality/multi-sensory interface arena. Here, advanced display techniques are often adopted only as a result of their "high-tech" qualities. However, when implemented in a tele-operation or telepresence system, those same qualities could then limit the utility of the system, because the human user spends more time trying to adapt to the limitations of the interface technology, rather than benefiting from the displayed content of the application itself.

To overcome this problem, the interactive facilities for UMV supervision and control must evolve from the early performance of a hierarchical task analysis (at least), and a domain-specific analysis (at best; e.g., [1]). The interactive facilities must also take into consideration the domain-specific issues covered by the task analysis (and the subject of Static/Dynamic Allocation of Function specifically (e.g., [2,3,4,5]) and must take into consideration the human factors limitations of candidate interface technologies (e.g., [1]).

6.3.2 Standards/Methodologies/Best Practice and Gaps

Academic and military laboratory-based RandD programmes have often failed to generate realistic interface design guidelines to support the human operator in controlling UMVs safely and efficiently. Furthermore, the international standards community also admits that, as far as advanced computer-based human interfaces are concerned, it is further behind in the delivery of underpinning standards for new technological developments, such as unmanned systems, than it would like to be.

There have been a few attempts historically to collate information of relevance to the hazardous environment tele-operation community. As well as a highly relevant Oak Ridge publication [6], UK research in the early 1980s led to the production of a Human Factors Design Handbook [7]. This handbook was specifically concerned with the design of human interfaces for Remotely Operated (submersible) Vehicles (ROVs) for the North Sea oil and gas industry and was used to design consoles for some vehicles still in service today. Whilst, today, this document is somewhat out of date (especially with regard to informational displays and the combination of video images with alphanumeric data or forms of symbology to assist in navigation), it still contains some information of relevance to general ergonomics practice, workplace and hand control design.

Another report that was produced with remote operations in mind focused on the needs of the nuclear industry and developments in support of mobile and manipulator systems for irradiated facilities [8]. Again, some of the information is dated. However, sections on hand control design and the application of recent (1990s) interactive technologies (e.g., from the virtual environments community) are still relevant.

More recently, organisations have been active in promoting the need to collate human factors information of relevance to the unmanned or remotely operated vehicle arena. A publication [9] was written to fulfil the role of an informative annex to pre-contractual and contractual documents relating to the design of future military or explosive ordnance disposal (EOD) tele-operated/robot systems.

For the foreseeable future, then, tele-operated vehicle and manipulator system designers will have to rely on general ergonomics and human factors texts, as adopted by the civilian and military human factors community. This is far from satisfactory because such texts tend to be outdated before they are even published. A good example here is the current generation of ISO standards pertaining to human-computer interaction, such as ISO 14915, which does not adequately address the needs of the growing multi-modal and synthetic environments communities. Furthermore, the refresh cycles for these documents are very long-term indeed. Some standards, such as DEF STAN 00-25 in the UK (*Human Factors for Designers of Equipment*) are only now being considered for updating (or "future proofing") – many of the contents dating back to, for example, [10].

Ageing and irrelevant contents of existing texts also force the authors of contemporary standards and guidelines to take short cuts when attempting to make their documents relevant to a particular domain or system. One notable example of this is Appendix B3 (Human Computer Interface) of STANAG 4586 – *Standard Interface of the Unmanned Control System (UCS) for NATO UAV Interoperability*. Here, the absence of any appropriately packaged knowledge relevant to UMV human interface design (and relevant knowledge certainly exists in the UMV community), has forced the authors to use unmodified extracts from ISO 9241 – *Ergonomic Requirements for Office Work with Visual Display Terminals*.

At a very high level, then, the human-system interfaces provided as part of an integrated UMV system, regardless of the environment on which that vehicle is to be deployed (land, air, sea surface, subsea, space, etc.) must, at their most basic level, be capable of:

- Presenting the operator with appropriate (video and/or synthetic) imagery from the remote environment to enable situation awareness, waypoint planning, real-time control (driving, flying, etc.), obstacle avoidance (including friendly forces or civilians), payload selection, payload operation and general manipulation to occur safely and efficiently;
- Presenting the operator with appropriate real-time data to enable him/her to transition seamlessly between supervisory and tele-operation modes of control without significant loss of situation awareness (encompassing the status of both the vehicle – power, performance, command link integrity, "health" and other subsystems – and the remote environment);
- Providing the operator with control and data input devices appropriate for remote control, manipulation, subsystem and payload selection/operation tasks; and
- Presenting the human operator with mission-specific data, adopting formats that avoid any cognitive conflict or mental resource monopolisation , thereby compromising real-time control or supervision.

These presentational requirements must not be hindered in any way by the selection and/or design of the human interface hardware, and displays and controls must support the human operator's intuitive exploitation of the information he/she is provided. More detailed design principles and guidelines, including reference to the development of UMV workstations and portable consoles (originally written in support of human interface design for unmanned EOD vehicles) can be found in [9].

Where appropriate, and in the absence of access to physical UMV assets or appropriate deployment scenarios, Virtual or Synthetic Environment (VE/SE) test beds should be given serious consideration in the quest to fill the human factors knowledge gaps in both interface technology provision and in the definition of the role of the human operator in supervising or directly controlling one or more UMV systems. Early examples of such test beds, based on low-cost approaches to simulation (e.g., using Microsoft DirectX, .NET or games engine technologies) are available. The *Alchemy* experimentation system, for example, developed as part of the UK Human Factors Integration Defence Technology Centre programme, supports Human Factors investigations of advanced display and control devices for UMV interaction, from head-mounted displays to video console-like hand controllers [11]. *Alchemy 2* (Figure 6-1), the latest version of this test bed, has evolved from a programming-intensive environment to become a highly reconfigurable "serious gaming" implementation and has already been demonstrated in the context of an iSTAR UAV (Allied Aerospace) and "marsupial" (SPAWAR) land vehicle deployment (tele-operated), together with a new man-portable, twin turbofan UAV concept (Kestrel Aerospace, UK).

Figure 6-1: Alchemy 2 Experimentation System.

6.3.3 References for Guidelines and Gaps

[1] Stone, R.J. (2004). "Rapid Assessment of Tasks and Context (RATaC) for Technology-Based Training". Proceedings of I/ITSEC 2004, Orlando.

[2] Rouse, W.B. (1977a). "Human Interaction with an Intelligent Computer in Multi-Task Situations". Proceedings of the 11th Annual Conference on Manual Control, NASA Ames Research Centre, pp. 130-143.

[3] Rouse, W.B. (1977b). "Human-Computer Interaction in Multitask Situations". IEEE Transactions on Systems, Man, and Cybernetics; 7(5), pp. 384-392.

[4] Rieger, C.A. and Greenstein, J.S. (1982). "The Allocation of Tasks between Human and Computer in Automated Systems". Proceedings of International Conference on Cybernetics and Society; Seattle, Washington, IEEE, 204-208.

[5] Milgram, P., Van De Graaf, R.C. and Wewerinke, P.H. (1985). "Control Loops With Human Operators in Space Operations – Part 1: Human Engineering Analysis, Synthesis and Evaluation Techniques". National Aerospace Laboratory, The Netherlands; NLR TR 84116 L, Part 1.

[6] Draper, J.V. (1985). "The Human-Machine Interface in Mobile Teleoperators". Proceedings of the Department of Energy Workshop on Requirements for Mobile Teleoperators for Radiological Emergency Response and Recovery (Foltman, A.J., Ed.). Dallas, TX: Argonne National Laboratory.

[7] Stone, R.J., Day, P.O., Rogers, H.G. and Kelly, C.J. (1984). "ROV Workstation Design Handbook". UK Department of Energy Contract No. E/5B/CON/753/802. Unpublished British Aerospace Document Ref. JS10014.

[8] Stone, R.J. (1992). "British Nuclear Fuels Limited Agreement No. A022595 (Schedule 2): Natural Man-Machine Interface Study Final Report, Volume 1 – Review and Requirements". Advanced Robotics Research Limited (ARRL) Report No. ARRL.92 002.

[9] Stone, R.J. (2002). "Human-System Interfaces for EOD Telerobots: Introduction to Human Factors Design Guidelines". Unpublished report prepared under contract to the QinetiQ Centre for Robotics and Machine Vision (Contract No. C4004-20747).

[10] Van Cott, H.P. and Kinkade, R.G. (Eds., 1972). Human Engineering Guide to Equipment Design. New York: McGraw-Hill Book Company.

[11] Stone, R.J. (2005). "Project Alchemy: An Experimental Human Factors Test Bed for iSTAR UAVs in Urban Operations". Proceedings of HCII 2005; Las Vegas.

6.4 DATA INPUT TECHNOLOGIES

6.4.1 Manual Input Devices (keyboards, mice, control sticks, touch devices, etc.)

6.4.1.1 Description of Technology

Controllers that involve a manual input include single-purpose buttons, levers, and joysticks as well as a keyboard for text entry and a mouse for pointing and selecting text and icons presented on a display.

Trackballs and touch pads can also perform much the same function as a mouse. Since these control devices are commonly used in a variety of applications, no further description will be given herein. A less frequently employed manual control device employs touch sensitive technology in which operators press the display surface over the appropriate label or icon presented on a display to make a selection. A directory of manufacturers of touch screen devices is available [1].

Touch sensitive displays can sense that the display is being touched by a pen or finger and can send this information, along with the touch location, to a host controlling system. A variety of technologies have been employed to implement touch screens: resistance, optical, capacitive, and acoustic techniques [2]. These technologies can be grouped into two classes: touch screens that use an overlay that responds to pressure and screens that are activated when the finger or pointer interrupts a signal. All implementations are considered robust and operation is simple and easy to learn. As a more "direct" selection input method, a mechanical intermediary (mouse) is not needed; nor is there need for positional feedback (cursor) consuming display space. Rather, the touch screen serves as both an input and output device, capitalizing on direct eye-hand coordination.

6.4.1.2 Actual or Potential Applications to UMVs

In existing UMV control stations, operators typically employ the manual control devices just described, with the exception of the touch screens. There is, however, potential UMV applications of touch screens, especially in future supervisory control stations that feature more automated control of vehicle and camera movement. Since the operator will have less dedicated hands-on (e.g., stick and throttle) control in these stations, the majority of operator interaction will involve monitor and control of subsystems. Thus, touch screens have a potential application to these UMV stations, as function selection using touch screens has proven to be fast and accurate, as well as easy to learn in other ground-based applications [3]. Moreover, since the number, shape, size, location, and label of the touch-sensitive fields are under software control (compared to electromechanical controls), they can be easily reconfigured corresponding to periodic upgrades made in the UMV system control architecture. It should be noted, though, that touch screens are only appropriate when inputs are limited and well defined [4]. Touch screens are not suited for tasks requiring precise positioning such as drawing and graphical input. Additional guidelines on the design of touch screens (touch-selection strategies, button size, feedback strategies, mouse-emulation strategies, touch biases screen angles [5] and numeric keypads [6]) are available for use in applying touch screens for UMV stations.

6.4.1.3 Technology Maturity, Challenges, and Unresolved Issues

In that manual control devices have historically been used in UMV ground control stations, technology maturity is not an issue. Even touch screens are commercially available and widely used in banks, restaurants, airports, and ground-based command/control stations. Thus, a Technology Readiness Level (TRL) Rating #6 can be assigned to touch screens, since this input device has been demonstrated in relevant environments. (Conventional manual devices are at a TRL Rating #9.) Even though touch screens and other manual devices are appropriate for UMV station applications, there still remain challenges in determining how best to employ these devices in future supervisory stations such that the controls support decision making or execution in a timely and error-free manner.

Dedicated (single-function) manual control devices need to be easily located, grasped, and manipulated. All the information and control devices needed for a particular set of activities should be in close proximity and ideally available with less than two key presses. Proper and consistent formats, abbreviations, symbol meaning, control assignments, procedures and rapid (< 0.2 seconds) feedback need to be employed so the action required and status of control operation is intuitive to the operator [7].

Human-engineered design of multi-function controls is also needed [8]. With multi-function controls, operators select menu options presented on a control station display using buttons on the periphery of the display or a touch sensitive screen. While a large number of systems can be controlled via one control device, there is a danger in the operator spending excessive time with the head down in order to progress through multiple interface layers or menus. This can lead to distraction from the primary tasks of maintaining situation awareness and supervising UMV control. Besides effective labelling so the function associated with each switch is obvious, navigation between the menus and the interactive sequences need to be efficient.

Function selection using touch screens presents additional challenges. Arm fatigue and screen obstruction by the hand may occur. This is especially true with pen input, along with the hassle of having to pick up the fragile pen for each input. Operators must be more attentive to visual and audio feedback due to the lack of kinesthetic feedback associated with manipulating real switches and knobs [9]. Given the size of an operator's finger/pointer and the parallax potential of the screen, selection of small targets or closely spaced functions is also difficult. There are also limits on the types of interactions that can be accomplished. For instance, selecting and dragging objects, inputting precise positioning, rubber band line drawing, pop up menu selections and other interactions for which the mouse is well adapted are not suited for touch screen operation. To enable these more complex interactions, the touch screen needs to sense the degree, or pressure, of contact or sense multiple points of contact, besides just sensing the push and release of touch. If the system isn't able to sense at least two levels of pressure, then auxiliary devices must be used for signalling [9]. One potential solution is to use different sensor densities which allow the creation of display areas with different resolutions [10].

6.4.1.4 References for Manual Input Devices

[1] Buxton, B. (2005). A directory of sources for input technologies. Available at: http://www.billbuxton.com/InputSources.html#anchor693936

[2] Bullinger, H.-J., Kern, P. and Braun, M. (1997). Controls. In: Salvendy (Ed), Handbook of Human Factors and Ergonomics, 2nd Edition, New York: John Wiley and Sons, Chapter 21, pp. 729-771.

[3] Sears, A. and Shneiderman, B. (1991). High precision touch screens: Design strategies and comparisons with a mouse, International Journal of Man-Machine Studies, 34, 593-613.

[4] Plaisant, C. and Sears, A. (1993). Touchscreen interfaces for alphanumeric data entry. In: B. Shneiderman, Ed., Sparks of Innovation in Human-Computer-Interaction. Norwood, NJ: Ablex.

[5] Scott, B.M. and Conzola, V. (1997). Designing touch screen numeric keypads: effects of finger size, key size, and key spacing. Proceedings of the Human Factors and Ergonomics Society, 360-364.

[6] Lewis, J.R. (1993). Literature review of touch-screen research from 1980 to 1992. IBM Technical Report 54.694. New York: IBM Corporation. Available at: http://drjim.0catch.com/touchtr.pdf#search=‘Beringer%20AND%20touch%20and%20input

[7] Lind, J.H. and Burge, C.G. (1992). Human Factors Problems for Aircrew-aircraft Interfaces: Where should we focus our efforts? In "Advanced Aircraft Interfaces: The Machine Side of the Man-Machine Interface", AGARD-CP-521, 3-1-3-12.

[8] Calhoun, G.L. and Herron, E.L. (1982). Pilot-machine Interface Considerations for Advanced Aircraft Avionics Systems, In: "Advanced Avionics and the Military Aircraft Man/Machine Interface", AGARD-CP-329, Chapter 24, 1-7.

[9] Buxton, W., Hill, R. and Rowley, P. (1987). Issues and Techniques in Touch-sensitive Tablet Input, Computer Graphic, 19(3), 215-224.

[10] Westerman, W., Elias, J.G. and Hedge, A. (2001). Multi-touch: A new tactile 2-D gesture interface for human computer interaction. Proceedings of the Human Factors and Ergonomics Society, 632-636.

6.4.2 Speech Recognition Systems

6.4.2.1 Description of Technology

Automatic Speech Recognition (ASR) technology, which transforms an operator's spoken words into machine text, has been around for over 50 years [1]. Investigating this technology for use in military systems has been ongoing for over 30 years. In the early 1970s, researchers quickly realized that along with the rapidly advancing computer technology came the equally rapid proliferation of knobs, dials, switches and other traditional interface devices needed by operators to control that technology. Not only would these conventional controls begin to take up large amounts of physical space, but their sheer numbers would also quickly overwhelm the human operator's ability to effectively manage them all, especially in time critical situations. For these reasons, researchers began to explore the possibilities of using ASR, also known as Direct Voice Input (DVI), in addition to speech synthesis technology, which translates text into machine-spoken words, to control and display information [2,3,4].

Much of the early research in the military use of speech recognition technology was applied to the single-seat aircraft cockpit [5,6,7,8,9] where pilots must not only use stick, throttle and rudders to keep the aircraft in the air, but must also manage an increasing number of sophisticated onboard computer systems for offensive, defensive, surveillance, and a myriad of other tasks. It was becoming increasingly clear that without another crewmember in the cockpit, the complexity of these systems would eventually overwhelm the single pilot's ability to effectively manage the system. If an accurate DVI system were used, the pilot could verbally ask the aircraft to accomplish simple tasks like changing a radio frequency or navigate to a new waypoint [10,11]; much the same way a pilot might ask his/her co-pilot to accomplish these tasks.

Since 1987, the United States Defense Advanced Research Projects Agency (DARPA) and the National Institute of Standards and Technology (NIST), have sponsored objective benchmark tests in the DVI research community [12]. Figure 6-2 shows how the state of the art of the technology has improved significantly in recent years, and as a result, is rapidly becoming widely accepted by users [13].

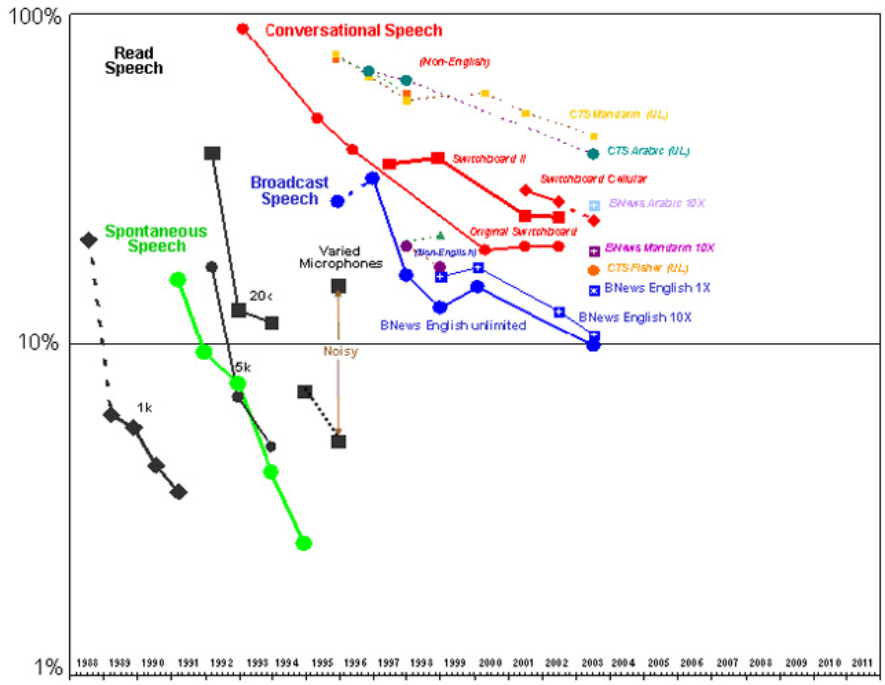


Figure 6-2: NIST Speech Recognition Benchmark Test History.

Speaker dependent, discrete word recognition systems requiring users to repeat all vocabulary words and phrases numerous times before ever using the system have given way to speaker independent, continuous speech systems requiring no prior system training. Small, static vocabularies of 300 or less words have been replaced by very large, dynamic vocabularies containing tens of thousands of words and are robust enough to handle both male and female native and non-native speakers. The systems have also matured to the point where they can achieve high recognition rates in very noisy environments [14,15]. As a result of these recent technology advances, numerous military applications for using this natural and intuitive communication method are being pursued [16,17,18,19,20].

6.4.2.2 Actual or Potential Applications to UMVs

Numerous laboratory and field studies have shown that the intelligent use of speech recognition and synthesis technology can significantly improve the operator interface. In an aircraft simulation experiment using DVI to control single-seat cockpit functions [10], speech was found to be a viable control method for changing radio frequencies and navigation coordinates and for designating ground targets in an air attack mission. Using this technology, data entry task performance was increased, while overall pilot workload was reduced. When multi-tasking, pilots were able to enter numeric data into the onboard systems while their hands and eyes were busy. In addition, the interaction of using speech technology combined with an automatic target cuing system resulted in superior performance in selecting and marking targets.

In a study designed to compare data entry performance using speech recognition with manual (keyboard, mouse, push button) methods [17], pilots performed data entry and query tasks while flying a simulated Unmanned Aerial Vehicle mission. The operators were asked to perform a continuous flight/navigation control task while responding to intermittent data entry, data query and emergency checklist tasks. The results showed that data entry task performance improved approximately 40% when pilots used DVI, in comparison

to performance with conventional manual controls. The operators also judged the speech system to be easier to use than manual controls and resulting in lower workload.

The U.S. Army has also investigated the use of speech recognition technology for ground vehicle operator interfaces in a M1A2 tank simulator [21]. Results showed that tank commanders could perform tasks (such as calls for fire, medical evacuation requests and radio setup) faster using speech input compared to using conventional push buttons and a joystick cursor controller. Operators could eliminate the menu "drill-down" search for the "call for fire" menu function by simply speaking a "call for fire" command. Operators also reported higher perceived levels of workload when using the conventional input method.

Quite a bit of research in the potential uses for speech recognition technology in helicopters has been done in the United States [22,23,24], United Kingdom [16] and Canada [25,26] including a successful flight test of an ITT system in the mid and late 1980s. Successful applications of speech technology in a number of ground based helicopter simulation platforms as well as a number of actual helicopter flight tests in the United Kingdom, Canada and the United States are described in [16,26].

Interesting uses for speech recognition technology for training systems have also been fielded by the U.S. Naval Air Warfare Center Training Systems Division (NAWCTSD) in Orlando, Florida. Some systems already in operational use include:

- A Carrier Air Traffic Control Center (CATCC) laboratory which realistically simulates the at-sea air traffic control environment. A speech recognition system replaces human pilot role players at the CATCC "C School", an advanced technical training facility in Pensacola, Florida that serves 90 Navy students per year plus team training for each of the Navy's carriers.
- A Tower Operated Training System (TOTS) laboratory, also in Pensacola, Florida, serves as a VFR Tower Air Traffic Control facility trainer that simulates the visual environment of a tower cab. TOTS features speech recognition/synthesis to enable direct student interaction with simulated pilots.
- The Navy's Advanced Radar Air Traffic Control (ARATC) trainer in Pensacola, Florida, realistically simulates shore-based air search and ATC radar systems using speech recognition/synthesis technology. The trainer provides for approach, arrival and precision radar final approach training by allowing students to communicate with simulated pilots.
- The Radar Air Traffic Control Labs (Radlabs) training system in Pensacola uses speech recognition to teach basic radar air traffic control procedures.
- The Amphibious Air Traffic Control Center (AATCC), a realistic trainer for the amphibious assault ship environment, uses speech recognition/synthesis to provide realistic interaction with the simulated pilots.
- A Landing Signal Officer (LSO) Trainer at Naval Air Station, Oceana, Virginia features a wide field-of-view projection of a carrier landing platform and allows the LSO student to communicate with the pseudo-pilot via a speech recognition/synthesis system.
- The Air Traffic Control Proficiency Trainer (APTS) is a table-top low cost trainer developed to train basic aircraft control and phraseology. APTS recently was delivered to 46 U.S. Navy and Marine Corps sites worldwide.

It is clear from the number of successful applications of this technology in both simulated and operational ground and flying platforms that speech recognition will be an important and integral component in future military operator interfaces. Indeed, operational use of the technology has already found its way into military

jet aircraft. In Europe's newest jet fighter aircraft, the Typhoon, DVI gives pilots the ability to control data entry functions such as changing radio frequency and displaying fuel status as well as more critical functions such as configuring radars and radios, all without having to move their hands from the flight controls. This DVI functionality serves as an alternative to using manual control methods when the pilot's hands and eyes are busy.

Following the Typhoon's lead, the new French Rafale aircraft will have a DVI capability, allowing the pilot to perform actions through spoken commands using a vocabulary of 90 to 300 words. The United States also plans to field a DVI capability in its new Joint Strike Fighter (JSF). Under a Memorandum of Agreement signed with the JSF prime contractor Lockheed Martin, Adacel (Melbourne, Australia-based) will develop a speech-enabled cockpit control system as part of the 10-year System Development and Demonstration phase of the JSF program.

Through the breadth of DVI research conducted since the 1950's, a number of "lessons learned" have emerged that serve, in part, as suggested guidelines for the successful implementation of this technology. These lessons, together with general guidelines published by Gardner-Bonneau, Rudnicky and others for telephony and other spoken language systems [13,27] should serve as a baseline set of guidelines for the successful implementation of speech recognition technology in military systems. Some of these include:

- Limit or eliminate speaker training – State of the art speech recognition systems no longer require speakers to "train" the system to recognize their voice commands. Modern DVI systems automatically adapt to the ambient acoustical environment and the nuances of a speaker's voice over time.
- Phrases that are from three words or longer are more distinctive and will be better recognized than short one or two word phrases. Military operators are trained to use short, terse speech on radios, but single-word commands are less distinctive and can be confused by the recognizer.
- Avoid commands that sound similar, but have different meanings. For example, "Turn backups on" and "Turn backups off" differ by only one phoneme and might be confused by the recognizer in a noisy environment.
- Define commands that are easy to remember, feel natural to say, and don't conflict with menu items or controls.
- Make use of "Macro" commands. Provide speech-recognition commands that add value by doing more than can be accomplished through a single click or keyboard equivalent.
- Use confirmation when necessary. For commands that could result in data loss, ask for confirmation before executing the command.
- Redundancy – Make every attempt to provide an alternative method for accomplishing an action. Don't make speech the only way for the user to accomplish a task.
- Errors will occur. Bias toward deletion errors, rather than errors of insertion or substitution.
- Have a clear understanding of the trade offs associated with using speech controls vs. manual controls and limit the use of speech for critical tasks that could be accomplished with a simple button push.
- Provide feedback to the user to indicate that the speech system either took or didn't take action.
- If graphic feedback is impractical, provide auditory feedback.
- At all times, the user should know what can be spoken.

- Do not overwhelm the user with feedback.
- A robust, continuous listening system is ultimately better than one that requires explicit user action. Unfortunately, in many military systems, due to high ambient background noise and/or other speakers in close proximity, a push-to-talk implementation is probably the best solution (push and hold a button to speak, release when done speaking).
- Offer an Undo capability. If Undo is not practical, use a confirmation protocol in cases where undoing a recognition error is troublesome.
- More complex applications need to track state, so that actions are interpreted in the proper context.
- Out-of-Phraseology Speech – Speech recognizers are not yet capable of understanding unconstrained human speech. Accordingly, applications are developed based on constrained vocabularies – but users often say words that are not in the legal vocabulary. Since the speech recognizer will try to make the best match, undetected out-of-phraseology speech could be processed with chaotic results. The challenge is to detect out-of-phraseology speech and reject it before it is post-processed.
- If a certain action is made available through speech, make sure that tasks involving this action can be completed via speech. If confirmation is required, make sure the user can speak the response (instead of forcing the user to type something, for example). Users prefer to stay in one input mode rather than switching back and forth depending on the task.

6.4.2.3 Technology Maturity, Challenges, and Unresolved Issues

Speech recognition and synthesis technology has already been fielded in military training and simulation systems using relatively small vocabularies. System prototypes have also been successfully demonstrated in operational environments such as aircraft cockpits and ground control stations, and are considered to be at TRL Rating #7. However, the ultimate goal of speech technology researchers is to create a system that listens, understands and behaves as well as (or better than) a human might in the same circumstances. Current state of the art systems are quite good at listening for predefined words and phrases, but are not yet very good at understanding unconstrained, natural language speech. Significant research in machine language understanding is currently being done for both commercial and military systems. The U.S. Defense Advanced Research Projects Agency (DARPA) has funded much of this work under their Communicator and other programs. Several leading universities have built variations of the Communicator architecture, including MIT's Galaxy Communicator system [28], Carnegie Mellon University's (CMU) Communicator system [29], and the University of Colorado's CU Communicator system [30].

6.4.2.4 References for Speech Technology

[1] Davis, K.H., Biddulph, R. and Balashek, S. (1952). Automatic Recognition of Spoken Digits, Journal of the Acoustical Society of America, 24 (6), 637-642.

[2] Beckett, P. (1987). Voice Control of Cockpit Systems. In: Information Management and Decision Making in Advanced Airborne Weapon Systems: Conference Proceedings of the Aerospace Medical Panel Symposium. Toronto, Canada, 27, 1-27.

[3] Werkowitz, E.B. (1984). Speech Recognition in the Tactical Environment: The AFTI/F-16 Voice Command Flight Test. In: Proceedings of Speech Tech '84 Voice Input/Output Applications Show and Conference, New York, NY.

[4] Lizza, G. and Goulet, C. (1986). Cockpit Natural Language. Proceedings of the 1986 National Aerospace and Electronics Conference. Dayton, OH: IEEE, 818-819.

[5] Barry, T., Liggett, K. and Williamson, D. (1997). Human-Electronic Crew Communication: Applications for Speech Recognition in the Cockpit. AFRL-HE-WP-TR-1999-0235. The Human-Electronic Crew: The Right Stuff? Tri-National Workshop Proceedings, Kreuth, Germany, 123-129.

[6] Gerlach, M. and Onken, R. (1993). A Dialogue Manager as an Interface between Aircraft Pilots and a Pilot Assistant System. In: Human Computer Interaction, Proceedings of Fifth International Conference on HCI (HCI International '93), Salvendy G. and Smith M.J. Eds., 98-103, Volume 2.

[7] Onken, R. (1995). Human-centered Cockpit Design Through the Knowledge-Based Cockpit Assistant CASSY, Paper presented at AGARD Symposium on Situational Awareness, DRG Panel 8, Technical Proceedings AC/243 (Panel 8), TP/7, AGARD/NATO Publication.

[8] South, A. (1997). Voice Recognition in Adverse Aircraft Cockpit Environments. In: Audio Effectiveness in Aviation, AGARD Conference Proceedings 596. AGARD-CP-596, p. 34-1.

[9] Steeneken, H. and Pijpers, E. (1997). Development and Performance of a Cockpit Control System Operated by Voice. In: Audio Effectiveness in Aviation, AGARD Conference Proceedings 596. AGARD-CP-596, pp. 33-1.

[10] Barbato, G. (1999). Lessons Learned: Integrating Voice Recognition and Automatic target Cueing Symbology for Fighter Attack Missions. Tenth International Symposium on Aviation Psychology, Columbus, OH.

[11] Williamson, D. and Barry, T. (1990). Cockpit Application of Voice Technology. Proceedings of the Eleventh Annual IEEE/AESS Dayton Chapter Symposium, IEEE, Dayton, OH, 57-60.

[12] Pallett, D.S. (2003). A Look at NIST's Benchmark ASR Tests: Past, Present, and Future, National Institute of Standards, http://www.nist.gov/speech/history/pdf/ NIST_benchmark_ASRtests_2003.pdf

[13] Gardner-Bonneau, D.J., Ed. (1999). Human factors and voice interactive systems. Boston, MA: Kluwer Academic Publishers.

[14] Williamson, D., Barry, T. and Liggett, K. (1996). Flight Test Performance Optimization of ITT VRS-1290 Speech Recognition System. In: Audio Effectiveness in Aviation: Conference Proceedings of the Aerospace Medical Panel Symposium, Copenhagen, Denmark, p. 35.

[15] Williamson, D., Barry, T. and Liggett, K. (1996). Flight Test Results of ITT-1290 in NASA OV-10. Proceedings of the 15th Annual International Voice Technologies Applications Conference (AVIOS '96), San Jose, CA, 335-345.

[16] Collins, P., McKelvy, T. and Abbott, M. (2004). Mature Speech Recognition Technology for Military Helicopters. Paper presented to the American Helicopter Society 60th Annual Forum, Baltimore, MD.

[17] Draper, M., Calhoun, G., Ruff, H., Williamson, D. and Barry, T. (2003). Manual versus Speech Input for Unmanned Aerial Vehicle Control Station Operations. Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, Denver, CO: HFES.

[18] Williamson, D. and Barry, T. (2001). Speech Recognition in the Joint Air Operations Center – A Human-Centered Approach. Proceedings of the 9th International Conference on Human-Computer Interaction, New Orleans, LA.

[19] Williamson, D. and Barry, T. (2000). The Design and Evaluation of a Speech Interface for Generation of Air Tasking Orders. Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting, San Diego, CA, 750-753.

[20] Williamson, D., Barry, T. and Draper, M. (2004). Commercial Speech Recognition Technology in the Military Domain: Results of Two Recent Research Efforts. Proceedings of the 23rd Annual Applied Voice Input/Output Society (AVIOS) Conference, San Francisco, CA.

[21] Gross, J., Ciappara, N., Smist, T. and Benson, P. (1998). Evaluating the M1A2 Tank Commander's Interface: The Battle of Input Devices. Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting. Chicago, IL: HFES.

[22] Holden, J.M., Stewart, S.J. and Vensko, G. (1987). Speech Recognition in a Helicopter Environment. Proceedings of Military Speech Tech '87. Arlington, VA: Media Dimensions, 147-152.

[23] Holden, J.M. (1989). Two Helicopter Flight Tests of Voice Command. Proceedings of Military Speech Tech '89. Arlington, VA: Media Dimensions.

[24] Simpson, C. (1990). Evaluation of Speech Recognizers for Use in Advanced Combat Helicopter Crew Station Research and Development. NASA Technical Memorandum 90-A-001, Moffett Field, CA: Ames Research Center.

[25] Farrell, P., Churchill, L., Wellwood, M. and Herdman, C. (2001). Integrating Direct Voice Input in a helicopter crew environment. Proceedings of the 11th International Symposium on Aviation Psychology, Columbus, OH: OSU.

[26] Swail, C. and Kobieski, R. (1997). Speech recognition in the Helicopter Cockpit Environment. American Helicopter Society 53rd Annual Forum, Virginia Beach, VA.

[27] Rudnicky, A. Speech Interface Guidelines. http://www.speech.cs.cmu.edu/air/ papers/SpInGuidelines/ SpInGuidelines.html

[28] Polifroni, J. and Seneff, S. (2000). Galaxy-II as an Architecture for Spoken Dialogue Evaluation, Proceedings of the Second International Conference on Language Resources and Evaluation (LREC), Athens, Greece.

[29] Rudnicky, A., Thayer, E., Constantinides, P., Tchou, C., Shern, R., Lenzo, K., Xu, W. and Oh, A. (1999). Creating natural dialogs in the Carnegie Mellon Communicator system. Proceedings of Eurospeech, 1531-1534.

[30] Pellom, B., Ward, W. and Pradhan, S. (2000). The CU Communicator: An Architecture for Dialogue Systems, ICSLP, Beijing.

6.4.3 Additional Alternative Input Devices

Control devices that require a manual input impose restrictions on the operator's posture at the control station. Moreover, often they require the operator's gaze be directed at the control device and keep the operator's hands busy. There are numerous alternative controls that have the potential of providing novel means of interacting with UMV systems (e.g., hands-free, head-up input). Speech recognition technology, described in detail above, is one such technology. Non-conventional controllers that do not require a direct mechanical linkage between the operator and the input device are desirable for supplementing manual control. These alternative controls use signals from the brain, muscles, voice, lips, head position, eye position, and gestures. Some of these alternative controls are not yet commercially available or configured to facilitate integration with software under development for supervisory ground control station applications. Plus, the control achieved with many hands-free devices can be described as rudimentary. Any UMV application must take into account the respective dimensionality, accuracy, speed and bandwidth of control afforded by each alternative input device. Of these alternative input devices, speech recognition is the only technology with sufficient maturity and reasonable cost to be considered a near-term candidate for UMV control station application (TRL Rating #7). Since speech recognition technology was described earlier, this section includes brief overviews of the other alternative input devices whose Technology Readiness Levels range from 1 (EEG-based control) to 4 (gesture and eye-based control).

6.4.3.1 Gesture-Based Control

Hand and body gestures are an important component of communication. Gesture-based control seeks to exploit this channel for system control. The head, hands and body can be localized in space using trackers, video techniques, gloves or suits. Trackers (mechanical, electromagnetic, ultrasonic, and optical) are devices that allow one to measure the position and orientation of a body part in space [1]. Video techniques use image processing in order to identify a specific body part and then reconstruct its position, orientation and posture from two-dimensional video images [2]. Gloves and suits allow one to measure the relative positions and angles of body components. A directory of manufacturers of gesture capture devices is available [3]. Perhaps the most highly developed for system interactions are devices that employ glove-based electronic input. For applications in which the body and hands are involved in other activities, gesture-based control for hands-free operation may best involve detecting defined movements or positions of the operator's face or lips. For instance, optical and ultrasonic sensing technologies have been used to monitor an operator's mouth movement [4]. Two candidate control approaches include processing lip movements to provide 'lip reading' and translating symbolic lip gestures into control inputs [5].

One challenge with gesture-based control is identifying the beginning and end points of a gesture [1]. Typical solutions require the operator to take a 'default' posture between gestures, which serves as an anchor for the system. Another problem is detecting whether a gesture is addressed to the recognition system or is a part of normal interpersonal communication. Creating an active zone in which gestures are effective is one solution. Given that a specific gesture can be reliably discriminated from other activity, there are three styles of input that can be made with gestures [1]. With "direct input", the features of the gesture generate kinematically similar actions in the task domain (e.g., one-to-one control of a robot hand). With "mapped input", features of the gesture are mapped in some logical fashion to actions in the task domain (e.g., number of raised fingers indicating a parameter level). For "symbolic inputs", features of the gesture are interpreted as commands to the system. Most uses of facial expressions would fall into the latter category.

A potential UMV control application is illustrated by a study that examined continuous cursor controllers to designate targets in a stereoscopic three-dimensional tactical map resident in a manned aircraft station simulator [6]. The tracking volume was remote from the actual map so that hand movements were actually

made in a space close to the aircraft controls rather than within the volume of the map. This hand volume was reduced in scale so that hand movements were small compared with the size of the map. Results showed that target designation was faster with an ultrasonic hand tracker compared to designation with a four-axis joystick and a voice control system.

6.4.3.2 Gaze-Based Control

Operators naturally look at objects that they want to manipulate or use. For applications in which the operator views a display during control operations, harnessing the direction of gaze promises to be a very natural and efficient control interface. With such control, the computer initiates a predefined action once it receives an input based on the operator's point-of-gaze [7]. For example, gaze-based control could be used to designate points of interest on a display showing the view from a UMV-mounted camera. In addition to command-and-control applications, eye gaze can be used to detect visual attention and adapt the application based on this information [8]. For UMV applications, an intelligent decision aiding system might adapt the display format or tailor active speech commands based on the operator's point of gaze.

Unless the head is stationary (or, with some systems, held within a small motion box), determining eye gaze point also involves tracking the head position/rotation. Predominant eye tracking techniques can be classified as electro-oculographic, scleral coil, and optical methods. The optical classification is the largest of these categories and several methods within this category are the most commonly used for gaze-based control (e.g., corneal reflection trackers) [9]. The most frequent problem with gaze-based control is avoiding what Jacob [10] dubbed the "Midas touch" problem, with commands activating where ever the operator gazes. Careful interface design is also necessary to ensure that the operator's head and/or eye movements during task completion are natural and not fatiguing. Rather, natural head and eye movements should be used to provide a direct pointing capability that can be combined with other control modalities to command activation. Plus, the control design needs to take into account the accuracy limits of the gaze measurement system and its operation under variable illumination environments [9].

6.4.3.3 Electromyographic (EMG)-Based Control

EMG-based control uses the electrical signals that accompany muscle contractions, rather than the movement produced by these contractions, for control. Electrodes positioned on the surface of the skin detect these electrical signals (e.g., produced during an operator's clenching of the jaw). Next, the signal's parameters are translated into a control input with processing. The simplest algorithm employs an on-off control based on the level or rate of change of EMG activity. For example, if muscle activity at one electrode site exceeds some threshold, a prosthetic hand opens. Above-threshold activity at another site causes the hand to close. Control can also be based on EMG patterns rather than on levels or rates of change [11].

EMG-base control is also a potential input device for computer systems. Research has shown that signals extracted from electrodes on the forehead can be used to control the movement of a cursor to track computer-generated targets, as well as provide a quick, accurate discrete on/off response [12]. Thus, for UMV applications, signals from electrodes positioned on the operator's face may provide an alternative hands-free, head-up binary input response. Continued development in EMG-based control is required to optimize the signals employed, assess the stability of the electrode contact over time, and minimize the effect of operator movement and external electrical activity on signal recordings [11].

6.4.3.4 Electroencephalographic (EEG)-Based Control

Electrodes positioned on the surface of the scalp can record signals that represent a summation of the electrical activity of the brain. Although much of the recorded EEG appears to be noise-like, it does contain specific rhythms and patterns that represent the synchronized activity of large groups of neurons. There are two general approaches for translating the electrical activity of the brain into a control signal. In one approach, EEG patterns are brought under conscious voluntary control with training and biofeedback. For example, voluntarily increasing the EEG activity in a specific frequency band above a threshold might be used to turn a switch on. Another approach harnesses involuntary, naturally occurring brain rhythms, patterns, and responses that correspond to human sensory processing, cognitive activity or motor control. For example, an operator might imagine moving their right hand to push a button. The computer would recognize the EEG pattern associated with this movement preparation and operate the right-hand button without further operator action. Although detection of these signals is easily accomplished with inexpensive components, optimization of this alternative control requires minimizing the time required for signal processing and developing easily donned electrodes [11,13,14].

EEG-based control has been demonstrated for several applications: cursor control, alphanumeric input, binary operation of a neuromuscular stimulator, and roll-axis control of a motion simulator [14,15]. To date, only rudimentary control has been achieved. However, the ability to make a system input, albeit crude, without lifting a finger, uttering a sound, or directing one's gaze, inspires the dream of someday enabling a genuine "thought-based interface".

6.4.3.5 Multi-Modal Input Design

For both manual and alternative controls, in addition to considering the general adequacy of the candidate control, the specific mapping of the input device to control tasks must be addressed. It is unlikely that a single input device will be adequate for all the control functions required in a particular application. A specific input device will be elegantly appropriate for some control tasks and clearly inappropriate for others. For example, speech-based control can be useful for a variety of control tasks, but use of a speech command to designate a position on a two-dimensional surface can be cumbersome. In contrast, gaze-based control is efficient for designating a position on a map display, but an auxiliary control modality is more useful for commanding the system to act on the designated location. This action command might be a voice utterance [16] or voluntary facial muscle activation [17]. The use of multiple control input devices capitalizes on using voluntary gaze direction to rapidly designate a position and a voice utterance or voluntary facial muscle activation (EMG) commands to quickly initiate an action. A multi-modal input method that combines speech and touch screen operation was also found to create a more operator-friendly interface [18]. Thus, task-controller mapping must take into account how best to increase overall functionality by using multiple input devices.

6.4.3.6 References for Additional Alternative Input Devices

[1] McMillan, G.R. and Calhoun, G.L. (2000). Gesture-based Control, International Encyclopedia of Ergonomics and Human Factors, New York: Taylor and Francis, Inc., Volume 1:237-239.

[2] Borghi, F., Lombardi, L. and Porta, M. (2005). Basic hand gesture recognition for human-computer communication. Proceedings of the 11th International Conference on Human-computer Interaction, CD-ROM.

[3] Buxton, B. (2005). A directory of sources for input technologies. Available at http://www.billbuxton.com/InputSources.html

[4] Hashimoto, M., Yonezawa, Y. and Itoh, K. (2005). An alternative GUI interface using air propagated ultrasonic waves and sound in mouth. Proceedings of the 11th International Conference on Human-computer Interaction, CD-ROM.

[5] McMillan, G.R., Eggleston, R.G. and Anderson, T.R. (1997). Nonconventional controls. In: Salvendy (Ed), Handbook of Human Factors and Ergonomics, 2nd edition, New York: John Wiley and Sons, pp. 729-771.

[6] Liggett, K.K., Reising, J.M., Beam, D.J. and Hartsock, D.C. (1993). The use of aiding techniques and continuous cursor controllers to designate targets in 3-D space. Proceedings of the Human Factors Society, 11-15.

[7] Fono, D. and Vertegaal, R. (2005). EyeWindows: evaluation of eye-controlled zooming windows for focus selection. Proceedings of Human Factors in Computing System (CHI' '05). New York: ACM Press.

[8] Hyrskykari, A., Majaranta, P. and Räihä, K.-J. (2005). From gaze control to attentive interfaces. Proceedings of the 11th International Conference on Human-computer Interaction, CD-ROM.

[9] Leger, A. and McMillan, G. [Chairs]. (1998). Eye-based control. In: Alternative Control Technologies, RTO-TR-7, 44-59.

[10] Jacob, R.J.K. (1991). The use of eye movements in human-computer interaction techniques: what you look at is what you get. ACM Transactions on Information Systems (TOIS), 9, 152-169.

[11] Leger, A. and McMillan, G. [Chairs]. (1998). Biopotential-based Control. In: Alternative Control Technologies, RTO-TR-7, 44-59.

[12] Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Roe, M.M., Lu, L.G., Haas, M.W., Dennis, L.B., Pick, H.L., Junker, A. and Berg, C.B. (1996). Brain-body actuated control: Assessment of an alternative control technology for virtual environments. Proceedings of the 1996 Image Conference, 225-232.

[13] Wolpaw, R.J., Birbaumer, N., McFarland, D.J., Pfurtscheller, G. and Vaughan, T.M. (2002). Brain-computer interfaces for communication and control. Clinical Neurophysiology, 113, 767-791.

[14] Vaughan, T.M., Heetderks, W.J., Trejo, L.J., Rymer, W.Z., Weinrich, M., Moore, M.M., Kübler, A., Dobkin, B.H., Birbaumer, N., Donchin, E., Wolpaw, E.W. and Wolpaw, J.R. (2003). Brain-computer interface technology: A Review of the Second International Meeting. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11, 94-109.

[15] Nasman, V.T., Calhoun, G.L. and McMillan, G.R. (1997). Brain-actuated control and head mounted displays. In: J. Melzer and K. Moffitt (Eds.), Head mounted displays: designing for the user. New York: McGraw-Hill, 285-312.

[16] Hatfield, F., Jenkins, E.A., Jennings, M.W. and Calhoun, G.L. (1996). Principles and guidelines for the design of eye/voice interaction dialogs. HICS Third Annual Symposium on Human Interaction with Complex Systems, 10-19.

[17] Surakka, V., Illi, M. and Isokoski, P. (2004). Gazing and frowning as a new technique for human-computer interaction. ACM Transactions on Applied Perception, 1, 40-56.

[18] Nakagawa, S., Zhang, J.X. and Chengcharoen, W. (1995). Multi-modal interface with speech and touch screen. Proceedings of the Sixth International Conference on Human-Computer Interaction, 213-218.

6.5 DATA DISPLAY TECHNOLOGIES

6.5.1 Head-Mounted Displays

6.5.1.1 Description of Technology

A head-mounted display (HMD) presents real video imagery or synthetically generated visual imagery via a head-mounted optic system with very small displays attached to a helmet, visor, or set of spectacles (see Figure 6-3). There is a wide range of technologies and approaches associated with today's HMD systems, including the type and quality of image display, monocular versus bi-ocular design, the ability to display colour images, the ability to effectively display stereoscopic three dimensional (3D) images, system size/weight, and the ability to concurrently view the local external environment. As this section is intended primarily as a short summary of the HMD technology and its relevance to UMVs, interested readers should refer to [1,2,3] for more comprehensive coverage of this area.

Figure 6-3: Examples of HMDs.

HMDs can support either immersive or augmented reality applications, depending upon the transparency of the head-mounted optics. Immersive HMDs require the user to view only the image presented via the HMD optic system, while augmented reality HMDs allow the user to "see-through" the HMD display, thus combining imagery from the HMD with the surrounding real-world visual field. This section only considers immersive HMD display technology, i.e., designs that occlude the subject's view of his/her immediately surrounding physical environment. Augmented reality technology is discussed in Section 6.5.2.

This section also only considers head-coupled HMD applications (i.e., the HMD visual image is updated in response to head movements, via a position tracking sensor that provides the computer with the current head location/orientation information). Head-coupled HMDs enable a synthetically generated visual scene to be continually modified in response to head movements so that, no matter how the user moves, objects in the viewed scene appear to remain in stable locations (thus providing the impression that the user is moving within the virtually generated space). Alternatively, in tele-operated robotic systems, the motion of the user's head can be used to control the position/orientation of a remote camera (or other robotic action) [4]. Head-coupled HMDs offer the highest potential degree of immersiveness and utility. However there remain many limitations due to existing technology, as will be discussed below.

Depending upon the particular UMV application, it may be critical for a HMD to convey accurate depth information in an intuitive and accurate manner. Although all HMDs can convey a sense of depth and distance using conventional two-dimensional depth cues (including linear perspective, interposition, relative size, texture gradients, etc.), certain HMDs also have the potential to portray depth via various stereoscopic techniques. Stereopsis can provide an intuitive, unambiguous cue for depth and it dominates most other depth cues. Dichoptic presentation involves using two monitors to portray a scene, one monitor per eye, each with its appropriate viewpoint for that eye's position in space [5]. This method utilizes binocular fusion to yield stereopsis. Electronic shutter glasses use one monitor to present a stereoscopic image by providing two alternating views of a scene (corresponding to the viewpoint disparity between the two eyes), synchronized to the frame rate, such that one interleaved frame in each pair is presented to each eye. Section 6.5.3 contains more information on various 3D display technologies.

6.5.1.2 Actual or Potential Application to UMVs

HMDs have been found to enhance wide-area search and intercept operations performed by manned aircraft [6]. A potential advantage of head-coupled control versus manual control over one's viewpoint is the addition of ecologically relevant proprioceptive cues which provide motion information based on vestibular inputs, joint angles, muscle lengths, and tendon tension during head movements. Head-coupled HMDs are also hypothesized to reduce cognitive processing demands in achieving new viewpoints. Some studies have suggested that head-coupled configurations facilitate awareness of areas already searched, thereby potentially reducing the re-scanning of those same areas [7]. Thus, HMD technology may benefit UMV operators, especially since reduced situation awareness is often a by-product of current UMV control station designs [8,9]. It is theorized that a HMD could enhance the operator's large-area searches and overall spatial orientation of the remote environment, while larger, high-resolution fixed console displays could be reserved for any target fine discrimination tasks. Other expected advantages of HMDs include hands-free control and intuitive operation. However, studies investigating the benefits of HMDs have so far produced mixed results. Below is a summary of the recent research in the area, categorized by type of UMV: UAV, UGV and UUV.

6.5.1.2.1 UAVs

HMDs have been demonstrated to have a positive impact of certain UAV control tasks. A study [10] explored UAV operator control of a remotely-operated helicopter using an omni-directional camera controlled by a head-coupled HMD. The HMD system was found to promote operator immersion, encouraging a feeling of 'presence' as though the operator was physically in the vehicle. The HMD also resulted in faster and more accurate completion times in a simulated helicopter control task, compared to the alternative of attempting to control the helicopter while viewing it directly from the ground. These results support claims that HMDs can provide increased situation awareness. However, the non-HMD condition was somewhat lacking in that it did not include a fixed-display out-the-window view from the helicopter, so it is unclear whether viewpoint location or head-coupled HMD provided the observed improvements. Another study [11] explored HMDs for small UAV applications. The task involved piloting a small UAV past a ground target and then turning around at various distances to re-acquire that same ground target. The researchers found that the use of a head-coupled HMD resulted in faster and more successful re-acquisition of the ground target than when using a conventional display of imagery from the UAV's nose-camera. However the horizontal field-of view was nearly twice as large for the HMD as compared to the conventional display, which may have contributed to these findings. Additionally, all subjects complained of discomfort when wearing the HMD. Nonetheless, there is research support for the proposition that HMDs can provide greater situation awareness resulting in increased UAV operator performance.

An early experiment at the TNO Research Institute was conducted to explore the applicability of a head–slaved camera system in UMV applications [12]. To overcome some possible drawbacks of HMDs (e.g., weight), a HMD was compared with a head–slaved dome projection (Figure 6-4). To overcome the possible drawbacks of transmission delay, a method was introduced to compensate for the spatial distortions. This technique, called delay handling, preserves the correct spatial relation between the viewing direction of the camera and the operator, by presenting incoming images in the camera viewing direction at the moment the images were recorded, and not in the actual viewing direction of the operator. The results showed that delay handling is successful in supporting the perception of correct spatial relations. No differences in task performance were found between the actual HMD and the dome projection.

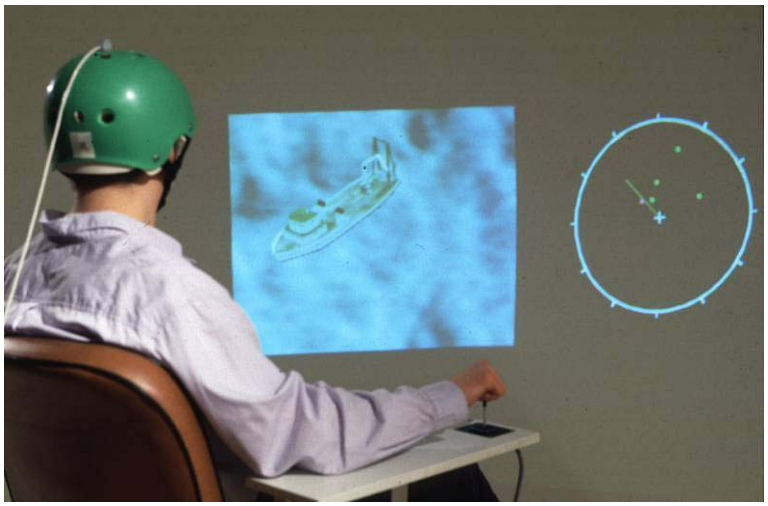


Figure 6-4: Dome Projection in which the Camera Direction is Head-Coupled, and the Operator Receives High Quality Proprioceptive Feedback on Camera Viewing Direction.

In follow-on studies at TNO, researchers compared operator performance with head–coupled camera control, and HMDs with manual camera control [13,14]. Subjects had to locate targets as fast as possible. The results showed that head-slaved camera control increased search speed, but enlarged the search path as compared to manual (joystick) control. An increased susceptibility, during head–slaved control, to mismatches between visual information and proprioceptive information may account for these findings. Additional measures of head movements showed that eye–head coordination was altered during head–slaved camera control. Since in these experiments, proprioceptive feedback was available in the manual control condition as well (the images were presented on a projection screen under the correct camera viewing direction), the findings imply an additional advantage of head–slaved control compared with manual control without proprioceptive feedback (as would be the case when using a fixed monitor). However, [15] found that employing simulated HMD images projected onto a large screen resulted in higher UAV operator performance than when they used an actual HMD, and [16] found that use of a conventional joystick for UAV control resulted in better performance than the head-coupled HMD. These latter results converge with several other studies, as detailed below.

Two experiments were conducted at the U.S. Air Force Research Laboratory to evaluate the usefulness of HMDs for UAV tasks involving the search for ground targets (Figure 6-5) [17,18]. The overall approach was to compare the utility of a manual joystick with associated stationary display monitor (the Baseline Condition) to that of various head-coupled HMD configurations. Specifically, gimbal camera orientation (azimuth and

elevation angle) was controlled via either a right-hand control stick or head-coupled HMD, while camera zoom was always controlled with the left-hand forward/aft stick. In one study [17], the task involved conducting a wide area search followed by a target identification task. The wide-area search was conducted by using the baseline configuration (control stick and stationary display monitor) or a head-coupled HMD. The target identification task was always conducted using the higher resolution stationary display monitor, as the HMD did not have the required resolution to afford fine discrimination. Thus, in the HMD conditions, there was a need to switch displays between search and identification tasks. The results failed to show any benefit for HMD-based configurations. Search time was shorter and workload was lower with the Baseline Condition than any of the HMD conditions. Additionally, many subjects experienced discomfort and simulator sickness symptoms with the HMD configuration.

Figure 6-5: UAV Workstation with Head-Coupled HMD and Stationary CRT Camera Displays.

A follow-on study [18] was conducted to specifically evaluate the utility of a head-coupled HMD for the SO's conduct of a 360-degree large area search for multiple ground targets. This study did not include the additional target identification step that had required a switch from HMD to a stationary display in the previous study. Six camera control/display configurations were evaluated; two involved the stationary display monitor (each with a different rate gain joystick) and four involved a HMD. The four HMD configurations varied in the degree to which the camera moved with head movements. One "hybrid" configuration was also evaluated whereby the gimballed camera orientation could be controlled with both the head and the joystick simultaneously. Results indicated fewer unique targets were prosecuted with the HMD than with the fixed display monitor. Head-coupled control also resulted in more duplicate target designations, higher workload, and lower situation awareness ratings. These results suggest that there is no clear advantage for head-coupled HMDs in the performance of large-area search tasks. In fact, performance significantly decreased in some experimental manipulations involving the HMD.

A similar set of studies were conducted utilizing a simulation of a smaller UAV [19,20]. One study compared a conventional display monitor to a HMD for target search, discrimination and designation tasks [19]. Although there were no differences between display conditions for target detection accuracy, the conventional display condition enabled more targets to be correctly identified from further away and allowed for more

accurate cursor designation of those targets. Additionally, subjects experienced far more discomfort (e.g., nausea, disorientation, eyestrain) with the HMD condition. In a follow-on study [20], these researchers explored the effect of including various auditory cues (mono, stereo, 3D spatialized) to the ground target location with the comparison of visual display conditions (conventional, head-coupled HMD). The results confirmed earlier findings that conventional displays resulted in significantly more precise target designations and fewer reports of discomfort. However, although HMD conditions yielded higher operator workload ratings than conventional displays across all conditions, 3D spatialized cueing reduced HMD workload levels significantly.

6.5.1.2.2 UGVs

HMDs have characteristics that potentially offer many advantages over conventional UGV operator control units [21]. Advantages identified included the system's light weight, decreased power consumption, daylight readability, and theorized improvements in operator situation awareness and telepresence. HMDs have also been demonstrated in UGV systems [22]. HMDs were found to be beneficial during a demonstration of the feasibility of utilizing a dune buggy as a UGV travelling complex terrain [23]. Other researchers conducted a study which found telepresence, created from use of stereo TV imagery displayed in an HMD, permitted operators' to drive UGVs at higher speeds and on steeper side slopes by providing an enhanced sense of spatial/geographic awareness [24]. There has also been implementation of HMDs into operational UGV systems. Man-portable UGVs are completing missions in current military operations with operators who wear monocular HMDs [25]. Soldiers are successfully controlling the UGVs with a portable joystick and HMD to explore cave complexes and suspected enemy compounds. Packbots' success in combat environments demonstrate HMDs' increasing and promising role in UGV control station design.

6.5.1.2.3 UUVs

Although few formal studies have been conducted in this area, the potential value of HMDs to UUV systems seems promising for underwater operations and operator training [26,27,37]. Other researchers have described the potential importance of providing UUV operators with meaningful cues for situation awareness, good workspace visibility, and vehicle behaviour feedback for effective performance [28]. The testbed they designed to address underwater telerobotics included a head-coupled HMD option. Though there is little research specifically addressing HMDs' effectiveness compared to other systems in UUV operation, the difficulties for operators controlling UAVs and UGVs are similar to those in underwater vehicles and so it is reasonable to assume that research on HMDs in these systems could transfer to UUVs.

6.5.1.3 Technology Maturity, Challenges, and Unresolved Issues

HMDs have improved considerably since they were first introduced into the commercial market. However, there is still much research needed before they achieve widespread appeal in military and consumer applications. Research areas discussed below include ergonomic issues, resolution, time latencies, field-of-view, and the occurrence of motion-sickness type symptoms.

6.5.1.3.1 Ergonomic Issues

Ergonomic issues associated with HMDs can be primarily attributed to anthropometric, biomechanic, and psychomotor concerns. Most HMDs involve some encumbrance by the user, though this varies with particular equipment chosen. Lack of fit is a primary complaint of users [29]. This includes inappropriate fit, movement limitations, excessive weight and/or size, and improper distribution of the weight. HMD weight

also has the potential to alter eye–head coordination. Suggestions exist for improving fit [29]. Newer displays are being developed to minimize size and weight such that they can be clipped onto existing eye-pieces, although other tradeoffs exist (limited resolution, small field-of-view, display placement within larger visual field, etc.). Furthermore, certain HMDs enable the display system to be removed or rotated out of the way to afford intermittent HMD use within a larger real-world work task. However it is unknown which method is most preferable for various UMV applications. Additional information regarding ergonomic issues can be found elsewhere [1,3,4]. Much research is needed to improve the many ergonomic issues of HMDs.

6.5.1.3.2 Spatial Resolution

Spatial resolution is a measure of the level of detail available in a visual display [2]. However, it can be a misrepresented parameter in HMD specifications. Often, resolution is described in terms of number of pixels in a display. However, the size of the display and its distance from the observer also contribute to the effective resolution. Increasing display field-of-view reduces effective resolution by enlarging each pixel in the same manner. Therefore a more effective manner in which to specify resolution is in terms of visual angle subtended per pixel, termed 'angular resolution'. Angular resolution is poor in most current HMDs, far lower than the resolving capability of the human eye (approximately 1 arc min visual angle or less [2]). Thus researchers have found resolution to be a limiting factor in HMD utilization [30]. An additional confusion with assessing spatial resolution of color HMDs is associated with the pixel-type used for determining angular resolution. Color display pixels are often formed by grouping 3 or 4 monocular pixels of different wavelengths (such as red, green, blue). Display manufacturers sometimes report the number of pixels and angular resolution based upon the total number of monocular pixels available instead of available color pixels.

Research is needed to better define spatial resolution requirements of HMDs for the range of envisioned UMV tasks. Additionally, research is needed to improve spatial resolution. One promising technology in this area is the virtual retinal display [31].

6.5.1.3.3 Time Latencies

Time delays exist between movements made by the user's head (which are tracked by a position-sensing device) and the response of the HMD scene to those movements, due to delays in position tracking and image generation [4]. Time delays between head movements and virtual image response result in loss of visual stability which can affect task performance and generate a sensory rearrangement between visual and vestibular cues of motion. These sensory rearrangements are believed to induce simulator sickness symptoms [32,33,34]. When the additional time delay associated with UMV datalink communications are factored in, the total delay can be on the order of several seconds. Additionally, time delays can affect user acceptance [7].

Specifications are needed for acceptable HMD time delay for various UMV applications, factoring in the delays associated with communication with the vehicle. Acceptable time delays for UUV operations may not be acceptable for fast-moving UAV systems. Research is also needed on methods to minimize the negative effects of time delay, such as through the use of prediction techniques for head motion.

6.5.1.3.4 Display Field-of-View (DFOV)

DFOV is the visual angle subtended by the display screen from a given observer location [35]. This parameter, described in terms of its horizontal and vertical components, is often desired to be large to promote a sense of immersiveness (i.e., presence) and to improve task performance through the utilization of

peripheral vision (limited DFOV displays result in the development of altered scanning strategies). However, a trade off that occurs when one tries to increase DFOV using an existing display is a corresponding reduction of screen resolution. Given a fixed display size, the only way to increase DFOV is to either magnify the display using optics or move the eye closer to the screen. In either case, pixel size increases in the same proportion as screen size (since both are fixed values). As pixel sizes increase, display resolution decreases. Research is needed to better understand DFOV requirements for various UMV applications. Additionally, the relation between DFOV and geometric field-of-view (zoomed in or zoomed out images) and its effect on UMV operator performance and comfort is needed [35].

6.5.1.3.5 Simulator Sickness/Cyber Sickness

Simulator sickness (also termed cyber sickness) is a form of motion sickness that occurs as a result of experiencing computer-simulated visual environments [35]. Symptoms include nausea, fatigue, headache, eye-strain, dizziness, malaise, and blurred vision. Besides the deleterious effects associated with simulator sickness, experiencing these symptoms may result in reduced desire to interact with the provoking system in the future, thus potentially hampering overall mission effectiveness. HMD usage has been strongly linked with increased levels of simulator sickness in many studies including those involving UMVs [17,19,20]. Although some guidelines exist, more research is needed to fully characterize and alleviate simulator sickness in UMV-related HMD applications.

6.5.1.3.6 Other Issues

Other research issues include the need to better understand and mitigate workload associated with HMD usage. Due to ergonomic concerns as well as the need to constantly move one's head to change one's viewpoint, workload and fatigue are real concerns associated with this technology [19] and mitigation techniques must be explored [18]. Head tracking technology and research is also needed to define minimum accuracy requirements for various UMV systems and to enable unencumbered operations [36]. Display brightness and contrast are also issues, especially for applications in outdoor environments.

6.5.1.4 References for Head-Mounted Displays

[1] Kalawsky, R.S. (1993). The Science of Virtual Reality and Virtual Environments, Reading, MA: Addison-Wesley.

[2] Kocian, D.F. and Task, H.L. (1995). Visually coupled systems hardware and the human interface. In: W. Barfield and T.A. Furness (Eds.), Virtual Environments and Advanced Interface Design, New York: Oxford University Press.

[3] Melzer, J.E. and Moffitt, K.W. (1997). Head-mounted Displays: Designing for the User. New York: McGraw-Hill.

[4] Durlach, N.I. and Mavor, A.S. (Eds.). (1995). Virtual Reality: Scientific and Technological Challenges, Washington, DC: National Academy Press.

[5] May, J.G. and Badcock, D.R. (2002). Vision and Virtual Environments. In: K.M. Stanney (Ed.) Handbook of Virtual Environments, Mahhaw, NJ: Lawrence Erlbaum Associates.

[6] Geiselman, E.E. and Osgood, R.K. (1994). Utility of Off-Boresight Helmet Mounted Symbology During a High Angle Airborne Target Acquisition Task. Proceedings of the SPIE Conference Helmet & Head-Mounted Displays & Symbology Design Requirements, Vol. 2218, 328-338.

[7] Pausch, R., Proffitt, D. and Williams, G. (1997, August). Quantifying Immersion in Virtual Reality. Proceedings of SIGGRAPH 97, Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Los Angeles, CA, 13-18.

[8] Gawron, V.J. (1998). Human factors issues in the development, evaluation, and operation of uninhabited aerial vehicles. AUVSI '98: Proceedings of the Association for Unmanned Vehicle Systems International, Huntsville, AL, 431-438.

[9] Gawron, V.J. and Draper, M.H. (2001). Human dimension of operating manned and unmanned air vehicles, NATO RTO Workshop on Architectures for the Integration of Manned and Unmanned Air Vehicles (SCI-100), Fairfax, VA.

[10] Koeda, M., Matsumoto, Y. and Ogasawara, T. (2002). Development of an immersive teleoperating system for unmanned helicopter. Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication, 25-27 September, 47-52.

[11] Naylor, M., Reid, L. and Delaurier, J. (2004). Investigation of the effect of head-slaved camera motion on image tracking in uninhabited air vehicles, Canadian Aeronautics and Space Journal; September, Vol. 50, 199-205.

[12] van Erp, J.B.F. and Kappé, B. (1996). Computer Generated Environment for steering a simulated unmanned aerial vehicle. TNO report TM-96-A039. Soesterberg, The Netherlands: TNO Human Factors Research Institute.

[13] van Erp, J.B.F. and van den Dobbelsteen, J.J. (1998a). Head–slaved and manual remote camera control with time delays. TNO report TM-1998-A076. Soesterberg, The Netherlands: TNO Human Factors Research Institute.

[14] van Erp, J.B.F. and van den Dobbelsteen, J.J. (1998b). Head slaved camera control, time delays, and situational awareness in UAV operation. Report TM-1998-A075. Soesterberg, The Netherlands: TNO Human Factors Research Institute.

[15] de Vries, S.C. and Padmos, P. (1997). Steering a simulated unmanned aerial vehicle using a head-slaved camera and HMD, Proceedings of the SPIE – The International Society for Optical Engineering; Vol. 3058, 24-33.

[16] de Vries, S.C. (2001). Head-slaved control versus joystick control of a remote camera, TNO-report TM-01-B008, Soesterberg, The Netherlands: TNO Human Factors Research Institute.

[17] Draper, M.H., Ruff, H.A. and LaFleur, T.C. (2001). The effects of camera control configuration on teleoperated target search tasks. Proceeding of the Human Factors and Ergonomics Society 45th Annual Meeting, 1872-1877.

[18] Draper, M.H., Ruff, H.A., Fontejon, J.V. and Napier, S. (2002). The effects of head-coupled control and head-mounted displays (HMDs) on large-area search tasks. Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, 2139-2143.

[19] Morphew, M.E., Shively, J.R. and Casey, D. (2004). Helmet-mounted displays for unmanned aerial vehicle control; Proceedings of SPIE – The International Society for Optical Engineering, Vol. 5442, 93-103.

[20] Dowell, S.R. and Shively, R.J. (2005). Synergy of virtual visual and auditory displays for UAV ground control stations. Proceedings of International Symposium on Aviation Psychology, Oklahoma City, 155-161.

[21] Browne, M. and Moffitt, K. (1996). A head-mounted display system for UGV control stations. AUVSI '96; Proceedings of the 23rd Annual Association for Unmanned Vehicle Systems International Symposium and Exhibition, Orlando, 15-19 July, 705-715.

[22] Kron, A. and Schmidt, G. (2005). Haptic telepresent control technology applied to disposal of explosive ordnances: principles and experimental results. Proceedings of IEEE ISIE 2005: International Symposium of Industrial Electronics, June 20-23, Dubrovnik, Croatia.

[23] McDonnell, J.R., Solorzano, M.R., Martin, S.W. and Umeda, A.Y. (1990). A head coupled sensor platform for teleoperated ground vehicles, Unmanned Systems, Fall, Vol. 8; 33-38.

[24] Metz, C.D., Everett, H.R. and Myers, S. (1992). Recent developments in tactical unmanned ground vehicles. Proceedings of The Association of Unmanned Systems International, Huntsville AL, 22-24 June.

[25] Hromadka, T.V. and Melzer, J.E. (2003). Results of using helmet-mounted displays to control robots in Afghanistan. Proceedings of SPIE – Helmet- and Head-Mounted Displays VIII: Technologies and Applications, Clarence E. Rash, Colin E. Reese, Editors, Vol. 5079, 222-231.

[26] Pioch, N.J., Roberts, B. and Zeltzer, D. (1997). A virtual environment for learning to pilot remotely operated vehicles. Proceedings of The International Conference on Virtual Systems and Multi Media VSMM '97, 10-12 September, 218-226.

[27] Boult, T. (2000). DOVE: Dolphin omni-directional video equipment. Proceedings of the IASTED International Conference on Robotics and Automation, August 14-16, Honolulu, Hawaii.

[28] Murray, S. and Murphy, D. (1996). Underwater telerobotics and virtual reality: A new technology partnership, PACON, Honolulu, HI, 20 June.

[29] McCauley Bell, P.R. (2002). Ergonomics in Virtual Environments. In: K.M. Stanney (Ed.) Handbook of Virtual Environments, Mahhaw, NJ: Lawrence Erlbaum Associates.

[30] Smith Sr., G.R. and Smith Jr., G.R. (2000). Virtual reality flight trainer for the UAV remote pilot. Proceedings of SPIE – The International Society for Optical Engineering, Vol. 4021, 224-233.

[31] Bayer, M.M. (2002). Retinal Scanning display: a novel HMD approach for Army aviation, Proceedings of SPIE – The International Society for Optical Engineering, Helmet and Head-mounted Displays VII, Vol. 4711, 202-213.

[32] Peli, E. (1995). Real vision and virtual reality. Optics and Photonics News, July, 28-34.

[33] Reason, J.T. and Brand, J.J. (1975). Motion Sickness. London: Academic Press.

[34] So, R.H. and Griffin, M.J. (1995). Head coupled virtual environment with display lag. In: K. Carr and R. England (Eds.), Simulated and Virtual Realities: Elements of Perception, London: Taylor & Francis.

[35] Draper, M.H., Virre, E.S., Furness, T.A. and Gawron, V.J. (2001). Effects of image scale and system time delay on simulator sickness within head-coupled virtual environments. Human Factors, 43(1), 129-146.

[36] Martinsen, G.L., Havig, P.R., Post, D.L., Reis, G.A. and Simpson M.A. (2002). Human factor requirements of helmet trackers for HMDs. Proceedings of SPIE – The International Society for Optical Engineering, Vol. 5079, 95-103.

[37] Stoker, B.P.H., Sims, M., Rasmussen, D., Hontalas, P., Fong, T.W., Steele, J., Barch, D., Andersen, D., Miles, E. and Nygren, E. (1994). The Application of Telepresence and Virtual Reality to Subsea Exploration, The 2nd Workshop on Mobile Robots for Subsea Environments, Proceedings ROV'94.

6.5.2 Augmented/Mixed Reality Technology

One reason that it is difficult to operate a remote UMV or its sensor is the "limited angular view associated with many remote vision platforms [which] creates a sense of trying to understand the environment through what remote observers often call a 'soda straw'" [1]. Another reason for the difficulty is the lack of the proprioceptive cues that would normally be available to a driver or pilot situated in a vehicle. One justification for the use of a head-coupled HMD as discussed in the previous section is to give the UMV/sensor operator a proprioceptive sense of the environment: the operator's proprioceptive sense of the relative pose of his or her head is used as a surrogate for the relative pose of a moving camera, albeit with mixed results.

An alternative approach is to represent multiple sources of sensor data in a mixed reality interface. Conventional interfaces do not integrate multiple sources of information into a coherent representation. For example, Figure 6-6 presents compass direction (upper left), laser range data (second from upper left), video (upper center), sonar data (upper right), map (bottom), and team information (right). All information required for robot control is available in the interface, but it is up to the operator to integrate this information into a meaningful and useful "picture" of what is going on. For operators, this means that the cognitive workload associated with just getting enough awareness to control locomotion can be high enough that it is difficult to interpret imagery, plan strategies, and do other higher-level tasks that are relevant for the mission. This problem is frequently solved by adding more humans to the team who are responsible for mission-level issues [2].

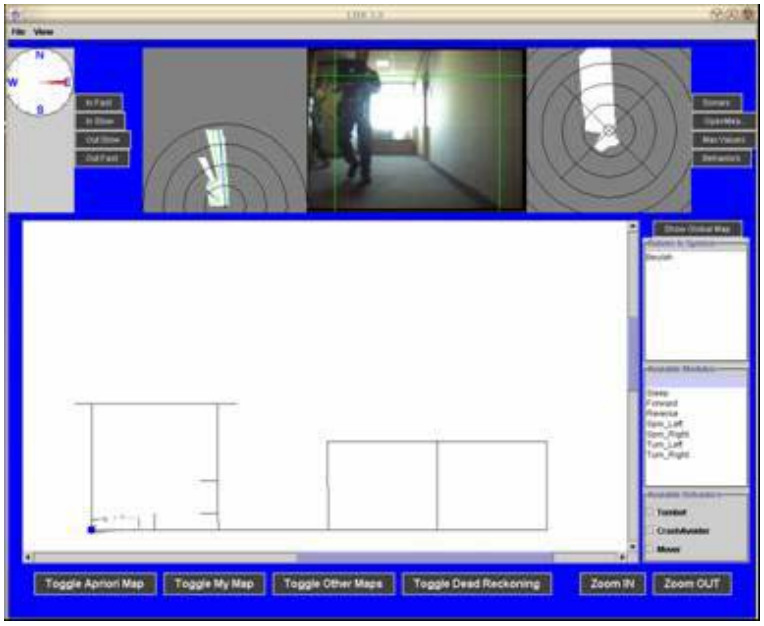


Figure 6-6: A Typical Interface Presents All Sensor Readings Side by Side.

6.5.2.1 Description of Technology

Another approach to this problem is to represent real sensor information in an ecological way [3]. In the robotic domain, one important and useful ecological technique is the use of mixed reality interfaces. Such interfaces combine real data with virtual elements, and range from augmented reality to augmented virtuality interfaces [4]. An example of such interfaces, taken from [5], is shown in Figure 6-7. In this interface, obstacles detected with a range sensor are represented by barrels in a virtual world. This virtual world is augmented with video from the robot. Fusing video, obstacles, and a representation of the robot allow the operator to perceive the relationship between the robot's "shoulders" and the obstacles in the world. This type of mixed reality display uses visualization techniques, such as those developed in gaming, as surrogates for the proprioceptive cues that are missing in remote operation.

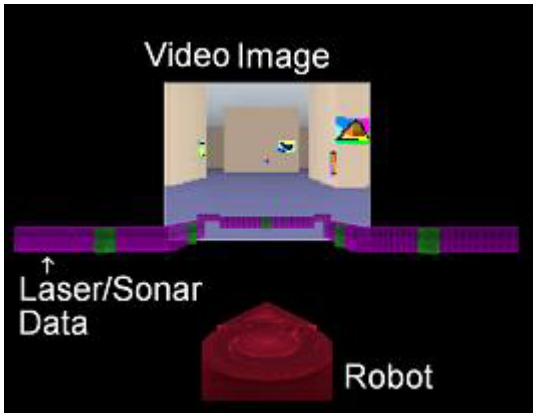


Figure 6-7: Integrated Display of Video, Range Readings, and Robot Representation.

There are a number of different ways for information to be presented in a mixed reality display. This section describes two techniques and shows examples from both UGV and UAV applications. The two techniques that will be shown are the chase perspective and the map-based perspective. Each technique is appropriate for certain types of tasks and modes of interaction between the human and a UMV. The idea that different visualization techniques have uses in different situations is a well-known result in display design, in general, and in augmented reality-based displays, in particular [6].

6.5.2.1.1 The Chase Perspective

The chase perspective is illustrated in Figure 6-7. This perspective presents sensor information in a way that supports locomotion, and is a typical representation used in racing games because it allows the direct perception of the relationship between the vehicle and the afforded directions of vehicle movement. Similar to the goal of using a head-coupled HMD to help an operator understand the pose of a movable camera, the chase perspective can be augmented with a visual representation of the pose of the camera relative to the robot. This is illustrated in Figure 6-8. Note, however, that large panning motions may require a shift from the chase perspective as illustrated in Figure 6-9.

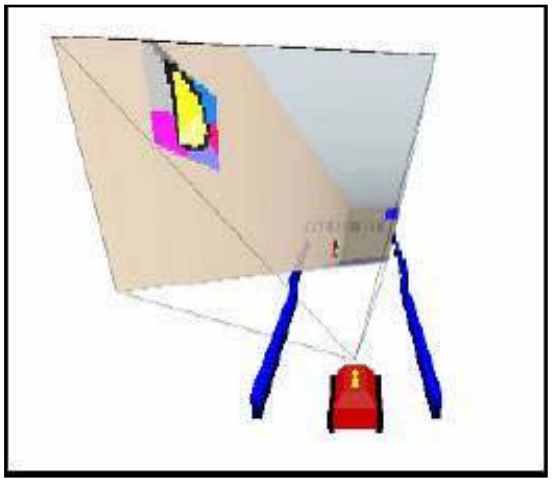


Figure 6-8: Representing the Pose of a Panning Camera.

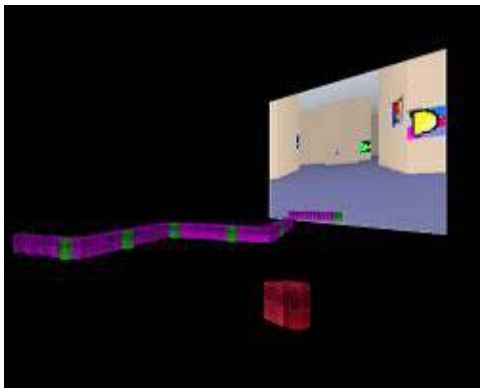


Figure 6-9: Depicting Camera Pose May Require a Perspective Change.

Similar visualization techniques can be used to represent other information such as deviations in terrain, the focal length of a zooming camera [7], and delays in receiving imagery from the robot via quickening [5].

A chase perspective can similarly be used to support aviation with UAVs. An example of the chase perspective is shown in Figure 6-10 [8,9]. In this display, a virtual UAV is included in the display to represent the pose of the UAV relative to the ground. This virtual UAV is overlaid on the video image received from the UAV and allows the operator to directly perceive the attitude of the aircraft with respect to the ground. (In the figure, two virtual UAVs are shown; one indicates the actual pose of the UAV as received from telemetry and the other indicates the commanded pose of the UAV.) The chase perspective shown in Figure 6-10 is taken from an interface that runs on a 5 inch or smaller display.

Figure 6-10: The Chase Perspective for a UAV.

Note that the chase perspective for the UAV is earth-centered rather than pilot-centered. When the operator is on the ground, banking right is not accompanied by a pilot-perceived change in the earth's horizon nor is it accompanied by other vestibular cues. Since the operator is on the ground, the chase perspective adopts a ground-based perspective wherein a bank command is depicted by having the virtual UAV dip its wing in the commanded direction.

It is possible to take this ground-centered perspective a step further. Since a fixed-mount camera rotates when the UAV banks, the operator must switch from a ground perspective to issue commands, to the UAV perspective to interpret video. This switch might be a cause of cognitive workload because the ground-based operator must interpret rotations in the video caused by a banking UAV. An interface that makes both video and bank angle have a ground-centered reference frame is shown in Figure 6-11. This interface is built for a control device called the PhyCon (for Physical Icon) [8]. Rather than using a handheld computer to issue commands to the UAV, a physical model of the UAV is used. When the operator banks the model, the actual UAV also banks. Although it is somewhat difficult to see in the figure, the pose of the aircraft is projected onto the video from a ground-centered reference frame using the chase perspective. This is a type of mixed reality interface [4]. (In practice, there are actually two virtual depictions of the UAV. The first virtual UAV is depicted using the actual telemetry from the UAV. The second virtual is depicted using the commanded pose from the PhyCon. Having both of these projected into the augmented reality display allows the operator to see that the actual UAV is responding appropriately to the commanded pose.)

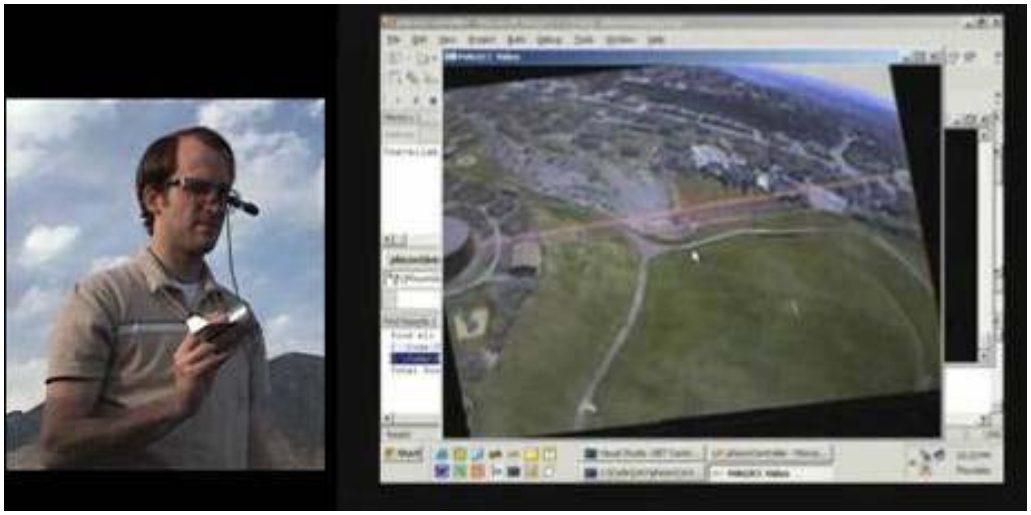


Figure 6-11: Rotating the Video Supports the Ground-Based Perspective of the Remote Operator.

The video feed is digitized and displayed on a computer (in this case, a laptop which presents the video through an eyeglass-mounted display). Prior to presenting the imagery, the telemetry from the UAV is used to rotate the image so that the horizon stays approximately level. This is depicted in Figure 6-11. Rotating the image so that the horizon stays level means that both the video and the UAV attitude are depicted in a ground-relative reference frame.

6.5.2.1.2 The Map-Based Perspective

The chase perspective primarily supports locomotion. When it is necessary to operate at the level of navigation or planning, it is often useful to have a map of some sort. For example, many potential UMV missions require some sort of exhaustive or heuristic search. Issuing commands for these searches and depicting the progress of these searches may be easier for the operator if a map-centered reference frame is used [6]. The task is to present map information in a useful way and then to integrate the video into the map using this map-centered perspective.

Figure 6-12 depicts a mixed reality display that integrates a virtual map with video from a robot [10,11]. The virtual map is a 3D rendering of a 2D occupancy grid map created using Konolige's map-building algorithms and software [12]. The video is depicted in this virtual world in such a way that video, map, and robot pose information are simultaneously visible. There are a number of desirable features of such an interface, including

a) The ability to determine what has been searched and what needs to be searched;

b) The ability to perceive how the robot is oriented with respect to landmarks in the world; and

c) The ability to augment map information with icons or other semantic labels.

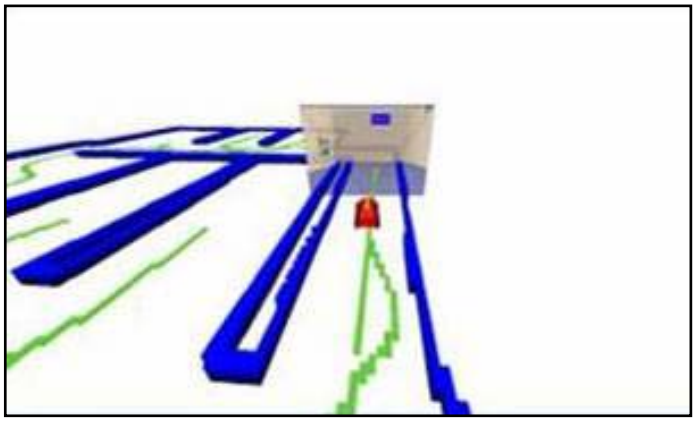


Figure 6-12: An Occupancy Grid Map Can be Used as the Basis for Navigation [10].

These same three elements, video, map, and robot (UAV) pose, can also be used to create a map-centered display for UAVs. In this case, the virtual world can be built on GIS terrain data or satellite imagery, the UAV pose can be depicted on this map with either a north-up or linked orientation [6], and the video from the UAV can be projected into the map with appropriate rotations so that the video is oriented in the same reference frame as the map. Some evidence suggests that using a map-based display helps operators more rapidly "cover" a particular region of interest by supporting better navigation [13]. The techniques described in the next section on various perspectives of the tactical space for 3D visualization can be simplified and applied to the augmented virtuality display. Such an interface is illustrated in Figure 6-13.

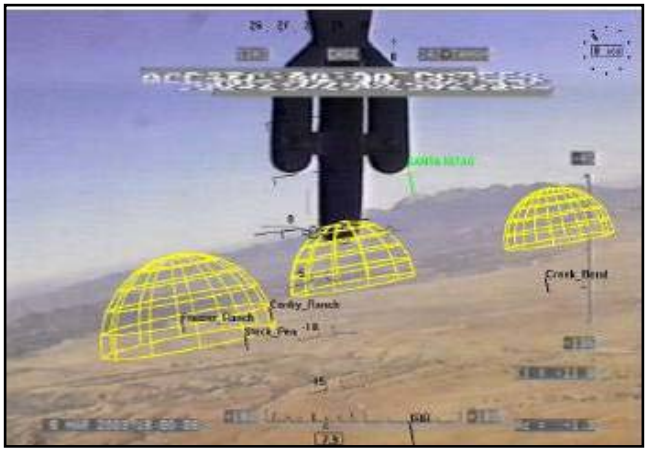


Figure 6-13: Augmenting a Terrain Map with Symbology Can Better Support Navigation and Sensor Management [23].

It is important to note that map-based interfaces have been used to construct augmented reality displays in aviation. These displays, which may be either heads-up or head-down and which may be retrofitted to older aircraft [14], are referred to as *synthetic vision* displays [15]. Several human factors studies have been conducted, many showing that there is an increase in navigation-related situation awareness with negligible loss in aviation-related situation awareness [16], presumably because a greater field-of-view and subsequent sense of realism can be obtained with such displays.

6.5.2.2 Actual or Potential Application to UMVs

Application of these display technologies to UMVs is an area of ongoing work. Research and development of mixed reality displays is being supported via Idaho National Laboratory under the Joint Robotics Program with the intent of developing a fieldable augmented virtuality display for UGVs in the very near future. As part of achieving appropriate levels of technology readiness, human factors studies have been conducted that provide strong evidence that the mixed reality interfaces described herein:

a) Make it easier to tele-operate a robot [5];

b) Make it easier to build a map of an area [11,13];

c) Work well with UGVs that have autonomy that allows interaction beyond tele-operation; and

d) Make it easier to use a panning camera [10].

The mixed reality displays for UAVs described herein run on small displays that may be appropriate for a dismounted control device of the type considered in the Future Combat Systems program. These interfaces have been demonstrated on laboratory UAVs that are equipped with the same automated controllers used on several class 1 military UAVs.

Research at the Air Force Research Laboratory's Human Effectiveness Directorate is exploring the value of mixed reality display concepts for UAV operation [17,23]. Research to date has focused on improving the situation awareness and performance of a UAV sensor operator for target search tasks through collaboration with Rapid Imaging Software, Inc. (see Figure 6-13). A recent study demonstrated a significant reduction in search time required to find ground landmarks when virtual marking flags were enabled. Future studies will develop guidelines for system update rate, symbology and labelling, declutter techniques and terrain depiction.

Like the occupancy grid-based displays for UGVs, terrain or image-based displays for UAVs are hypothesized to support better navigation, especially when navigation is complex or is performed under adverse visibility conditions [15]. However, an argument made in [18] indicates that the emphasis on realism in these displays may produce designs that are attractive but less effective than they should be. The authors call for designers to avoid "naïve realism" by using caricatures, icons, symbology and other abstracted representations of the kinds of information that an operator desires [19]. Added to this caution is the observation that overuse of symbology can create cluttered displays [20] and that certain types of disruptions can be very difficult to recover from even in the presence of clear symbolic labels [21]. Importantly, if information is sufficient to support precise navigation via augmented reality (synthetic vision) displays, it may also be sufficient to perform autonomous navigation. This may be especially important in areas such as search, where it is desirable to ensure that imagery from the camera efficiently and completely covers a region of the ground [22] and where screen size or team size is small enough that it is challenging for an operator to simultaneously aviate, navigate, and analyze imagery. The complexity of such navigation is illustrated in Figure 6-14.

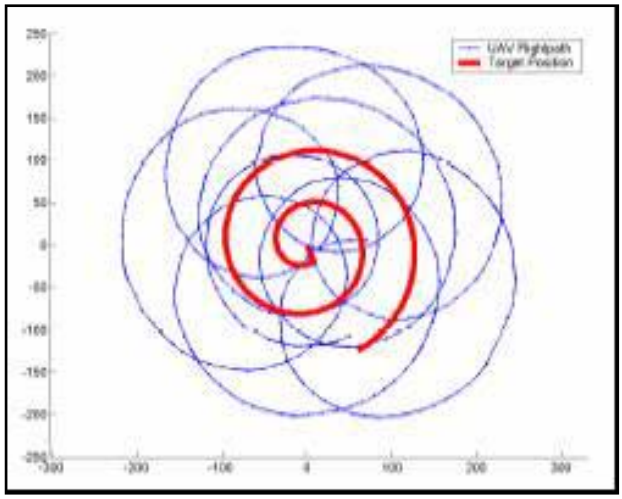


Figure 6-14: UAV Locations (thin blue lines) Required to Support Low Ground Speed, Complete Coverage by a Camera Footprint Spiral (bold red line).

Perhaps the most important potential application of mixed reality displays of the type described herein is in forming a common operating concept for many types of UMVs. The displays rely heavily on visualization techniques used in graphics and gaming, and therefore have some basis in being useful for a large class of operators in a large variety of situations. Moreover, the mixed reality displays described for UAVs are similar in concept to synthetic vision systems.

6.5.2.3 Technology Maturity, Challenges, and Unresolved Issues

The example interfaces described in this section are in the alpha or early beta stages of development. They have all been tested on physical platforms and there have been tests that confirm their usefulness in limited problems. A number of challenges and unresolved issues remain. The most pressing concern is the effects of terrain on visualization and control for both the UAV and UGV platforms. UGVs must frequently operate in outdoor, unstructured environments. While there is mounting evidence that in worlds with a level ground plane, such as inside a building, the UGV mixed reality interfaces work well, it is an open area how to adapt these interfaces to uneven terrain. A similar challenge exists with the UAVs. For class 1 UAVs, the operational altitude is frequently very close to the ground. Since the chase perspective and map perspective can make it difficult to convey information about height above ground to the operator, is may be desirable to use a terrain model and height above ground sensor to autonomously maintain a consistent height above ground.

Another very important challenge is to identify how to transition from one display perspective (e.g., chase view) to another perspective (e.g., map-view) when the operator shifts from one mission phase to another. Questions such as whether the display should automatically adapt to such shifts or whether the operator should explicitly command display changes are important to answer. Fortunately, there is a considerable literature on mode confusion, adaptive displays, automation management policies, and so on.

Other important challenges remain. These challenges include the following:

- Understanding the effects of confounding factors such as screen size, communications delay, and sunlight readability.

- Learning how to create virtual representations of shape shifting robots (e.g., the PackBot) or robots with manipulators.
- Identifying whether robot health information should be displayed as part of the virtual UMV representation, or as a separate part of the display.
- Learning how to represent the intent of the UMV's autonomy.
- Coordinating the mixed reality displays with other display concepts such as those required to support multiple robots, semantic mark-up of the map, symbology, and display declutter.

6.5.2.4 References for Augmented/Mixed Reality Technology

[1] Woods, D.D., Tittle, J., Feil, M. and Roesler, A. (2004, May). Envisioning human-robot coordination in future operations. IEEE Transactions on Systems, Man, and Cybernetics, Part C, 34(2), pp. 210-218.

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

[3] Vicente, K.J., Christoffersen, K. and Pereklita, A. (1995, April). Supporting Operator Problem Solving Through Ecological Interface Design, IEEE Transactions on Systems, Man, and Cybernetics, 25(4), 529-545.

[4] Milgram, P., Drascic, D., Grodski, J.J., Restogi, A., Zhai, S. and Zhou, C. (1995, February). Merging real and virtual worlds. Proceedings of IMAGINA '95.

[5] Ricks, B.W., Nielsen, C.W. and Goodrich, M.A. (2004). Ecological displays for robot interaction: A new perspective. International Conference on Intelligent Robots and Systems IEEE/RSJ, Sendai, Japan.

[6] Plumlee, M. and Ware, C. (2003). An evaluation method for linking 3D views. Proceedings of the International Conference on Coordinated and Multiple Views in Exploratory Visualization.

[7] Goodrich, M.A., Rupper, R.J. and Nielsen, C.W. (August, 2005). Representing Head, Shoulders, Eyes and Toes in Augmented Virtuality Interfaces for Mobile Robots, To appear in Proceedings of ROMAN, Nashville, TN.

[8] Quigley, M., Goodrich, M.A. and Beard, R.W. (2004, October). Semi-Autonomous Human-UAV Interfaces for Fixed-Wing Mini-UAVs. Proceedings of IROS 2004, Sendai, Japan.

[9] Cooper, J.L. and Goodrich, M.A. (2005). Portable Mini-UAV Control Interfaces for Non-Pilots, Presented at the 2005 Human Factors of UAVs Workshop, May 25-26, 2006. Mesa, AZ.

[10] Nielsen, C.W., Goodrich, M.A. and Rupper, R.J. (2005, August). Towards Facilitating the Use of a Pan-Tilt Camera on a Mobile Robot, To appear in Proceedings of ROMAN, Nashville, TN.

[11] Pacis, E.B., Everett, H.R., Farrington, N., Sights, B., Kramer, T., Thompson, M., Bruemmer, D. and Few, D. (2005). Transitioning unmanned ground vehicle research technologies. SPIE Proceedings, 5804: Unmanned Ground Vehicle Technology VII.

[12] Konolige, K. (2004). Large scale Map Building. Proceedings of the National Conference on AI, San Jose, CA.

[13] Bruemmer, D.J., Marble, J.L., Few, D.A., Boring, R.L., Walton, M.C. and Nielsen, C.W. (2005, July). Let Rover Take Over: A Study of Mixed-Initiative Control for Remote Robotic Search and Detection, In: IEEE Transactions on Systems, Man, and Cybernetics Part-A: Special Issue on Human-Robot Interaction, 494-504.

[14] Prinzel III, L.J., Comstock Jr., J.R., Glaab, L.J., Kramer, L.J. and Arthur, J.J. (2004). The Efficacy of Head-Down and Head-Up Synthetic Vision Display Concepts for Retro- and Forward-Fit of Commercial Aircraft. The International Journal of Aviation Psychology, 14(1), 53-77.

[15] Schnell, T., Kwon, Y., Merchant, S. and Etherington, T. (2004). Improved Flight Technical Performance in Flight Decks Equipped with Synthetic Vision Information System Displays. The International Journal of Aviation Psychology, 14(1), 79-102.

[16] Alexander, A.L. and Wickens, C.D. (2005). Synthetic Vision Systems: Flightpath Tracking, Situation Awareness, and Visual Scanning in an Integrated Hazard Display. In: Proceedings of the 13th International Symposium on Aviation Psychology. Oklahoma Cit, OK.

[17] Draper, M.H., Nelson, W.T., Abernathy, M. and Calhoun, G. (2004). Synthetic vision overlay for improving UAV operations, Proceedings of the Association of Unmanned Vehicle Systems International Annual Symposium 2004, Baltimore MD.

[18] Smallman, H.S. and St. John, M. (2005). Naïve Realism: Misplaced Faith in the Utility of Realistic Displays. Ergonomics in Design, Volume 13, Number 3.

[19] CHEX (Change History Explicit): New HCI Concepts for Change Awareness. In: Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society, 528-532, Santa Monica, CA.

[20] St. John, M., Manes, D.I., Smallman, H.S., Feher, B.A. and Morrison, J.G. (2004). Heuristic Automation for Decluttering Tactical Displays. In: Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, New Orleans, LA.

[21] Goodrich, M.A., Quigley, M. and Cosenzo, K. (2005). Task Switching and Multi-Robot Teams. Proceedings of the Third International Multi-Robot Systems Workshop, March 14-16, Washington, DC.

[22] Quigley, M., Barber, B., Griffiths, S. and Goodrich, M.A. (2005). Towards Real-World Searching with Fixed-Wing Mini-UAVs. Proceedings of IROS 2005. August 2-6, Edmonton, Alberta, Canada.

[23] Calhoun, G.L., Draper, M.H., Abernathy, M.F., Patzek, M. and Delgado, F. (2005). Synthetic Vision System for Improving Unmanned Aerial Vehicle Operator Situation Awareness. Proceedings of SPIE Vol. 5802, In: Enhanced and Synthetic Vision 2005, edited by Jacques G. Verly.

6.5.3 3D Visual Displays

Displayed information should fit the nature of the optimal mental processing operations required to perform the task. Wickens and Andre [1], Haskell and Wickens [2], and Wickens and Carswell [3] propose an interaction between the type of task performed and the type of display most suited for that task.

UAV supervisory control is an integrated task since the supervisor/operator must understand and combine location, angular, and rate of change information in 3D. Because 3D displays show the necessary information within a single spatial representation, it is proven that operator performance benefits more from these integrated 3D displays than from displays representing this spatial information in separate dimensions.

In this respect, we distinguish two types of 3D displays:

- 3D perspective displays. Perspective information may be considered as a subset of three-dimensional information (i.e., a stereoscopic image without binocular disparity).
- 3D stereoscopic displays. Stereoscopic displays are displays that create a true sense of depth.

6.5.3.1 Description of Technology

6.5.3.1.1 3D Perspective Displays

The graphical representation of the external world can be shown from the position of the observer, *egocentric* display information, or from a position somewhere in space, *exocentric* display information [4]). An egocentric presentation shows the external world only in one direction, suited to local guidance tasks, i.e., following a planned navigational track with limited preview. For example, when considering a flight task, pilots perceive themselves flying through the environment as seen from an ego-referenced frame [5]). This means that the display direction, left or right, always corresponds with the control direction. Exocentric displays separate the observer's eye point and actual position, showing the external world that surrounds the observer, thus assisting with global awareness. They represent the world either in a fixed geographical co-ordinate system (world-referenced; e.g., north up) or with respect to one's momentary position and orientation (ego-referenced; e.g., track up or heading up).

Research using egocentric perspective displays mainly has examined the navigation accuracy during local guidance tasks [6,7,8,9]. Research using exocentric perspective displays mainly focussed on world-referenced aspects: How effective will these displays support the situation awareness in a geographical environment? One study [10]) investigated the flight accuracy and orientation of pilots using two-dimensional (2D) plan-view, and 3D perspective north up, track up and heading up situation displays. Other investigations have examined pilots' perceptions of the geographical environment [11] and the assessment of collision risk [12]. Results have revealed that world-referenced exocentric displays can increase pilots' geographical orientation, but can hamper pilots' tracking performance in local guidance tasks because of the required mental transformations. For example, a north up display may cause confusion in an aircraft heading south [2,10]. Note that these investigations used more or less static scenarios. Ego-referenced exocentric display information, supporting the orientation of objects in space relative to the observer's momentary position, has hardly been a topic of much research activity. However, one may suppose that knowledge about the position of surrounding objects in space is of major importance.

Other research [13,14] has addressed methods of presenting perspective information on a display, investigating factors that influence the judgment of spatial information: grid-surface density, Geometric Field Of View (GFOV), Station Point Distance (STP), and target distance. It appeared that a perspective graphical presentation of the airspace provides a more natural (ecological) and compatible representation than a conventional plan-view display [15], but it was found difficult to estimate the exact position of computer-generated objects in that space. It is necessary, though, to carefully select the design parameters of the spatial information. For example, incorrect combination of GFOV and STP causes deformation of the presented image which leads to overestimation of the elevation angle [16].

Another important factor that affects the interpretation of perspective display information is scene dynamics during motion – the relative movement of the graphical components. In this respect, [17] distinguished motion references in avionic systems: with *inside-out* or ego-centered motion reference, the horizon rotates according to roll and pitch whereas the aircraft symbol remains stationary; and with *outside-in* or earth-centered motion reference, the horizon is stationary whereas the aircraft symbol rotates. Inside-out motion reference is compatible with the motion of the environment as it is observed from the cockpit; the aircraft is the reference, and the world is rotating. In contrast, outside-in motion reference shows the movements of the aircraft in a stationary world, representing the aircraft as a dynamic element in the real world. Research concerning the graphical representation of perspective display information as well as motion reference is reported here.

van Breda and Veltman [18] investigated 3D perspective displays in a simulated flight task. An aircraft guidance task was chosen with the following instruction: Perform a target acquisition task with a fighter aircraft, i.e., first locate a target that appears and then perform target interception (point the aircraft's nose toward the target as quickly as possible). For this task, it is of vital importance that pilots have correct estimations of the target's position, and of the route to be followed toward the target.

Egocentric displays are less suitable for this task because they provide only a limited view of the airspace in the flight direction. For high-quality task performance it is essential that the pilot obtain full preview, that is, being able to fully perceive the course of the stimulus (forcing function). Exocentric displays may meet this requirement, because targets beside or even behind are shown [19]. Both the display types were used in an experiment: an egocentric heading-up display for the initial aircraft following task (Figure 6-15) and an exocentric radar display for the interception task (Figure 6-16). Of the latter display, two types were investigated: a plan-view 2-D radar display, and a perspective 3D radar display. In the 2D display type, radar information consisted of an augmented circular plan-view display. The display centre represented the pilot's aircraft; a target symbol indicated the target position relative to the pilot's aircraft. The display was augmented with colour coding for relative target position, and a separate scale for target elevation was used. In the 3D condition, radar information consisted of an exocentric perspective spherical display, depicting the surrounding airspace as a dot pattern. Pilot's performance was analysed and evaluated in terms of target acquisition time, tracking performance and mental workload.

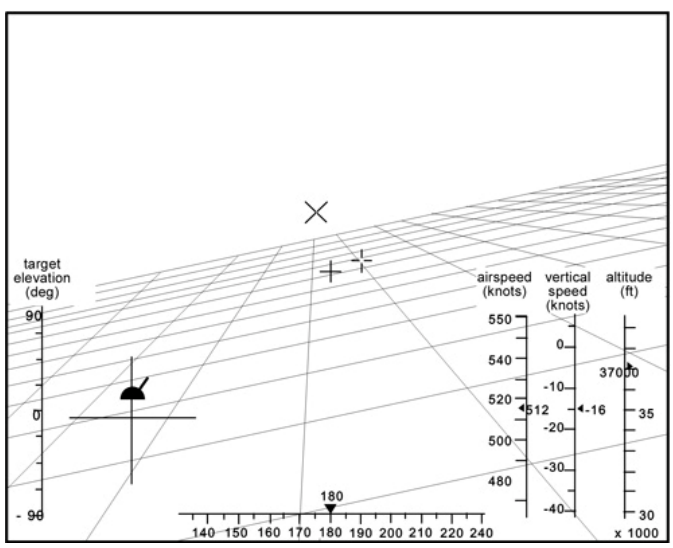
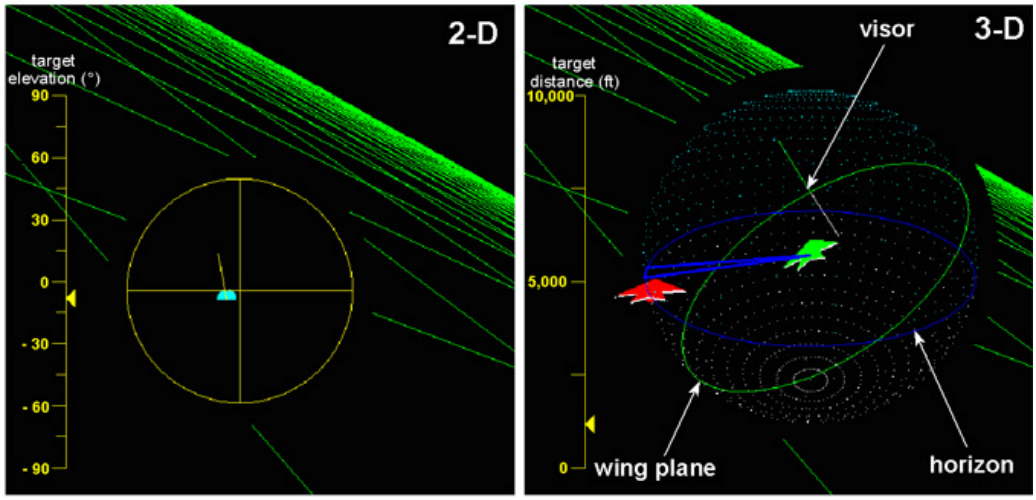


Figure 6-15: The Main Display.



a b

Figure 6-16: Overview of the Investigated Display Types.

The earth is represented by a grid, the target by an Xsymbol, the flight direction by a + symbol, and the length axis of the intercepting-aircraft by an _ symbol. Additional indicators for airspeed, vertical speed, altitude, and heading are shown. The radar display is presented in the lower left, in this case a two-dimensional plan-view radar image with an additional target elevation indicator.

The left figure shows a 2-D radar display with a separate target elevation indicator; the right figure shows a 3D radar display with a separate target distance indicator. The intercepting-aircraft symbol is always presented in the display centre. The line ahead, perpendicular to the wing plane, is the visor. Two 3D configurations were investigated: outside-in motion reference (the sphere with horizon and target symbol remain horizontal, whereas the intercepting-aircraft symbol rotates as a function of pitch and roll) and inside-out motion reference (the intercepting-aircraft symbol with wing plane remains horizontal, whereas the sphere with horizon and target symbol rotates as a function of pitch and roll). This figure shows a 3D radar display with outside-in motion reference.

The experimental results showed strong benefits of perspective displays for situation awareness support in the target acquisition task. A considerable reduction in the target acquisition time was obtained when pilots used a perspective radar display instead of a conventional plan-view display in the cockpit (Figure 6-17).

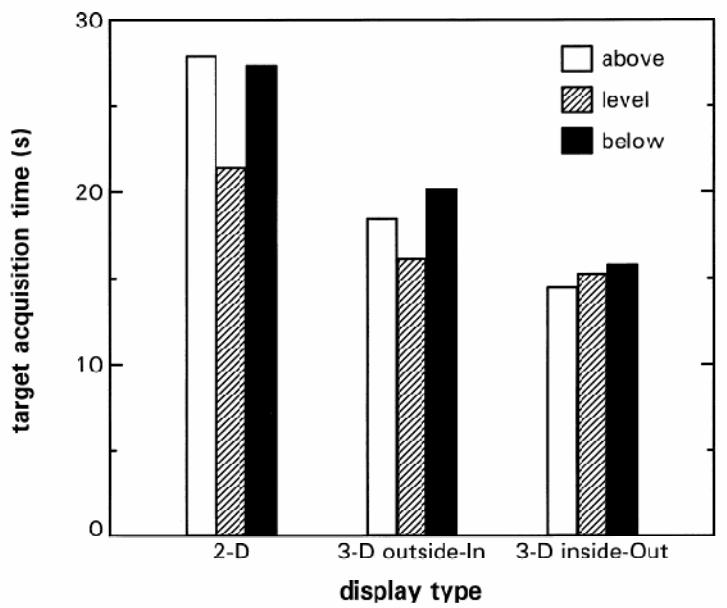


Figure 6-17: Target Acquisition Time as a Function of Two Dimensional (2-D) and Three Dimensional (3D) Display Type and Initial Position of the Target Aircraft, Averaged Across Participants.

This finding confirms one of the most important benefits of 3D perspective display representation as was observed by Wickens et al. [5]: The elimination of the time-consuming scanning that is necessary to go back and forth between the several parts of a display. In the 2-D display, a circular radar image indicating target azimuth and a separate linear indicator for the target elevation had to be scanned and mentally combined for target position estimation. In both the 3D perspective displays, target azimuth and elevation were presented in a single object. In the 3D perspective sphere, both the target elevation and target azimuth were integrated. For target acquisition, these variables represent information of close mental proximity.

Inside-out motion reference provided a direct relationship between the displayed movement of the scene and the perceived movements of surrounding objects. The display elements representing the outside world – three in this case – the sphere (globe) of the perspective radar display, the displayed horizon in the main (local guidance) display, and the visual horizon as seen from the cockpit, were consistent in this display, making the presented information ecologically valid [15,20]. The tracking task could therefore be considered as a natural process. The perspective sphere was presented by a dot pattern, providing adequate preview for tracking [21]. As was observed by [4]), 3D displays can be used very efficiently for local guidance tasks: in the current experiment the target acquisition time was reduced by more than *40%*. It is obvious that this is a major improvement in performance.

The subjective effort scores showed almost the same pattern as the performance data. The pilots felt that less effort was needed when perspective displays were used; in particular, they felt more comfortable with the inside-out motion display.

The strongest motive to ‘go three-dimensional’ is the inability to combine the two dimensional ‘bird’s-eye view’ with a graphical presentation of altitude and depth information. As a consequence, in all current systems altitude and depth information is presented as numerical read-outs. Representations lacking integrated altitude and attitude information complicate situation assessment in two ways [22]. Data that are difficult to acquire

are more difficult to use in making a decision. Also, without immediately evident altitude information, a decision maker may substitute arbitrary or situation-biased altitudes, that may be difficult to supplant even when the actual data are presented. With more realistic images of the environment and tracks in a 3D perspective (see Figure 6-18), decision makers argue, the interface becomes more natural and less effort is required to comprehend the current situation. It eliminates the burden of integrating and interpreting of multiple representations, abstract symbols, and textual read-outs. Some earlier experimental results with perspective displays confirm these expectations.

Figure 6-18: Example of a 3D Perspective Display (Adapted from Denehy, Johns Hopkins APL).

With the potential to improve performance, 3D perspective or stereoscopic displays, however, can also have their drawbacks. Inherent to the perspective view, objects are presented larger or smaller as a function of the operators viewing distance, location and angle. Objects close to the operator will be shown with much more resolution than objects at larger viewing distances. In many cases these differences will not necessarily reflect differences in tactical relevance and meaning. In their aspiration to design more realistic or natural representations of the environment researchers and software engineers also prefer to apply the principle of immersion to create the feeling of being part of that environment. Becoming embedded in the situation, however, can have some serious disadvantages. What you see becomes dependent on your own orientation. You have to look around not only to see what is happening miles away in front of you, but also what's happening directly behind your back. More realistic does not necessarily mean more functional.

One approach used at TNO is to let the environment become a 3D object in itself, such as a transparent cube that can be viewed from the outside and easily manipulated to see the environment from different perspectives and in different scales. In Figure 6-19, two pictures are shown that give an impression of this concept of a 3D tactical space. Simple geometric 2D and 3D symbols, comparable to the familiar naval tactical display symbology, are used to represent tracks.

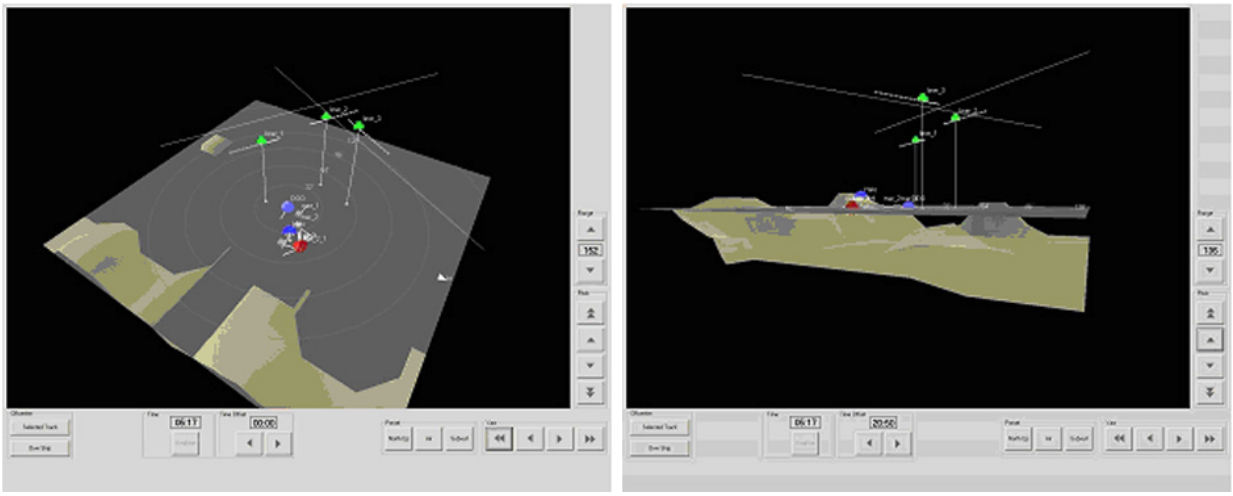


Figure 6-19: Examples of the Different Perspectives from which the Tactical Space Can be Viewed.

With one's own ship in the center of the tactical space as default, the operator has independent controls for horizontal and vertical zoom-in and zoom-out, to change the range or ceiling for which contacts are displayed. Size of symbols and track labels is held constant, irrespective of the zooming factor. With an off-center function, other tracks than the own ship can be selected to be presented in the center of the tactical space. By turning around and tilting the cube, the operator can view the environment from almost every perspective, be it air, surface or even subsurface, underwater. Independent of the selected perspective, all symbols and track labels remain in a steady orientation towards the user to guarantee good legibility. With a time-offset function the user can both playback or consult preceding moments in the situation and extrapolate the situation towards a future point in time. And dependent of platform type and armament, critical points in sensor and weapon coverage can be shown.

One of the most important functions of the tactical space is to have an environment that integrates information from different sources such as intel, traffic management, sensor and geographic information in a tactically meaningful way to support full situation awareness and assessment, especially when situations become less predictable or more complex as, for example, along the seashore. The concept of the tactical space, however, does not mean that all other ways of information display become superfluous, as confirmed in research done for the U.S. Navy [23]. There is no single display that can offer all information needed in an optimal way. The secret to good information presentation often lies in the diversity of multiple views on a situation and different graphical formats. With the tactical space well suited for the higher level tactical assessment, the 2D bird's eye view for instance, with its fixed orientation, more readily suits the fast localization of a track.

6.5.3.1.2 3D Stereoscopic Displays

Simply stated, 3D technology adds the sense of depth by imitating one or more of the visual depth cues. Here we describe how the 3D technologies achieve this result. A good overview can be found at the following website: http://www.stereo3d.com/3dhome.htm. More detail on the human factors aspects can be found elsewhere [24].

6.5.3.1.2.1 Convergence

Convergence can be activated by presenting (slightly) different images to the left and right eyes. The left-right difference is called the stereoscopic disparity. The most common methods are shutter glasses, polarized glasses, Red/Green glasses, and Head-mounted displays. They share the common disadvantage of constraining the user. Most importantly, eye contact is disturbed, hampering communication with others. These devices are therefore not well suited for group activities.

To avoid the constraints imposed by wearing optics in front of the eyes, so-called "auto-stereoscopic 3D displays" are being developed. The word "auto" signifies that the user does not need to wear an optical device. The optics are incorporated in the display, splitting the image in a left eye and a right eye component at the display instead. These optics are typically called "lenticular screens", and are glued to the flat-panel display. A lenticular screen consists of small lenses that bend the light from different display pixels in different directions. An inherent feature of auto-stereoscopic displays is that the head needs to be positioned at the right place. If for example the right eye is shifted 6 cm to the left, it will see the left eye image. Though solutions exist that allow some freedom of head movement, a price is paid in terms of an increase in cross-talk which reduces the viewing comfort [25] or in terms of added complexity in the form of a head tracker [26]. For a comparison of 3D methodologies on visual comfort see [27]. The disadvantages of the four main techniques can be summarized as follows:

- Shutter glasses: low luminance, flicker in daylight environments
- Polarised glasses: need to keep the head straight up
- Red/Green glasses: no color vision, chromatic aberration, cross-talk
- HMDs: image moves with the head, cables or weight

In the next section, a fifth technique (transparency) will be described.

6.5.3.1.2.2 Pictorial Depth Cues

Displays that contain symbols are oftentimes incompatible with the use of monocular depth cues because these cues tend to interfere with the clarity and standardization of the symbols.

6.5.3.1.2.3 Accommodation and Parallax

The 3D displays described above simulate the convergence depth cue, but do not provide accommodation and parallax, which means that the depth percept is incomplete. Parallax can be added by tracking the head movements and adjusting the view point accordingly. However, even with a fairly powerful computer a time delay between head movement and image adjustment remains noticeable. Except for one prototype 3D display in Oxford [28], the accommodation cue can only be added by imaging the scene at physically different distances. The most advanced system is the U.S. Navy sponsored "volumetric display" which achieves the accommodation cue by imaging on a rotating drum [29]. However, its large volume (approximately 1 m^3) makes it unsuitable for the type of applications we have in mind.

A relatively simple way to include accommodation and parallax in the depth percept is to optically superimpose two or more image slices through the scene. Such a transparent display presents "true depth" in the sense that all depth cues are present except for occlusion. An example transparent display is shown in Figure 6-20. A New Zealand Company [30] is the first to have marketed a compact transparent, 2-plane display. The display consists of two LCD filters, one placed in front of the other, making a subtractive

transparent display. Recently a 20-plane transparent display [31] has come on the market. We believe these developments to be highly significant, as argued in the next section.

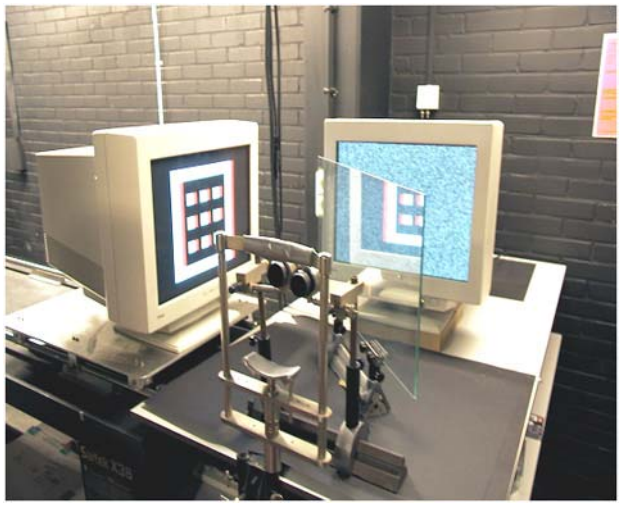


Figure 6-20: The Experimental Transparent 3D Setup at TNO Human Factors.

In Figure 6-20, two images are combined with a half-silvered mirror. Because the light from the two displays adds up, we call it an additive transparent display. The experiments examine the influence of various design parameters like amount of depth and scene content.

6.5.3.1.3 Transparent Depth Displays

6.5.3.1.3.1 Limited Number of Depth Planes

So far not much research has been done on transparent displays. This is probably because the limited number of depth planes makes them unsuitable for the display of 3D pictures and videos. The technologies described above can in principle present as many depth planes as the number of pixels in the display. We believe however that for professional applications involving the display of symbols, 2 or 3 depth planes will be sufficient to provide a large operational advantage. By analogy, many of the "full colour" cockpit displays by no means fully exploit their colour gamut; often a display only contains the four primary colours. Similarly the information content of control displays often times can be naturally divided into two (friend and foe) or three (above, below, and on the surface) layers. We therefore argue that the advantages of a transparent depth display will out-weigh the disadvantage of the limited number of depth planes.

6.5.3.1.3.2 Optimal Viewing Comfort and Depth Perception

In the case of transparent depth displays the depth percept is truly extra. The user does not pay a price in terms of resolution, colour, viewing angle, the need for special glasses, luminance, or viewing comfort as is the case with all other 3D displays. Secondly, our present research shows that the depth percept "pops-out" immediately while the other types of 3D displays require some amount of time for the depth percept to build up. Figure 6-21 shows how large the perceptual time delay is for unfavorable 3D stimuli. Thirdly, thanks to the parallax, occlusion of one object by another can easily be eliminated by moving the head sideways or vertically. This is important if two objects are located at the same x, y coordinates but different heights.

We therefore believe transparent planes to be very promising for the representation of computer generated data and symbols.

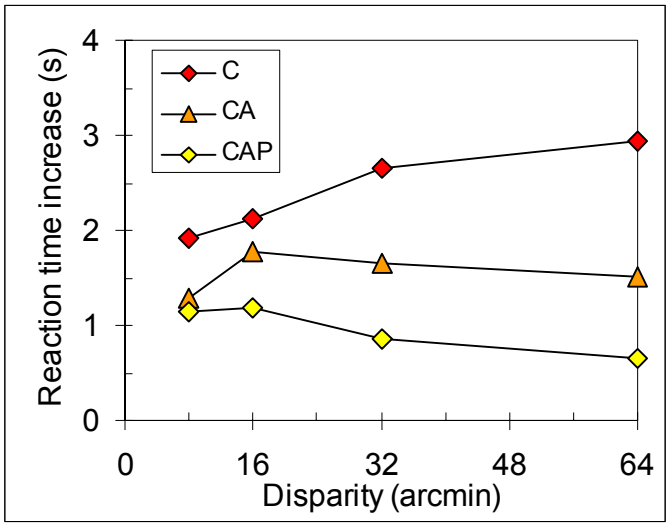


Figure 6-21: Data Substantiating the Claim that Accommodation (A) and Motion Parallax (P) Substantially Influences the Ease of Depth Perception when the Depth Gradient is Large.

Figure 6-21 shows the extra time required to perceive the depth relationship of two adjacent dots when a distracting object is added at a different depth. The horizontal axis contains the amount of depth difference. The increase in reaction time caused by the distractor is 1 to 2 seconds greater for the common type of 3D displays (C: Convergence cue only) than for transparent depth displays (CAP: accommodation and parallax as well as the convergence depth cue). These results imply that transparent depth displays are more natural to view than the standard 3D displays described in Chapter 3 and particularly suited for cluttered 3D imagery.

6.5.3.1.3.3 Subtractive and Additive Transparency

Another topic of current research at TNO Human Factors is design of the image content. We have shown that a transparent depth display, if designed wrong, can be very hard to fuse [32]. Secondly, the content of an additive transparent display (the front plane adds light: Figure 6-21) needs to be designed differently than a subtractive transparent display (the front plane removes light: e.g., Deepvideo [30]). The back layer of an additive display needs to be mostly dark, of a subtractive display mostly white. Otherwise the information in the other layer will not show up.

6.5.3.1.3.4 Occlusion

The 3D displays listed above do fully show occlusion, the phenomenon that the front object 'hides' the back object. An additive 3D display however is not able to show occlusion and a subtractive 3D display only partially. In addition, colour mixing leads to erroneous mixed colours: the colour of the front object is influenced by the object behind and vice versa. For example, a yellow object on a purple background will appear as red on a subtractive 3D display (Figure 6-22). For any application involving warning symbols this is unacceptable and is at the expense of unwanted colour mixing. We therefore argue that only an additive +

subtractive transparent 3D display will support all the depth cues listed in Section 6.5.3.1.2. This may therefore well be the next major step forward in 3D technology.

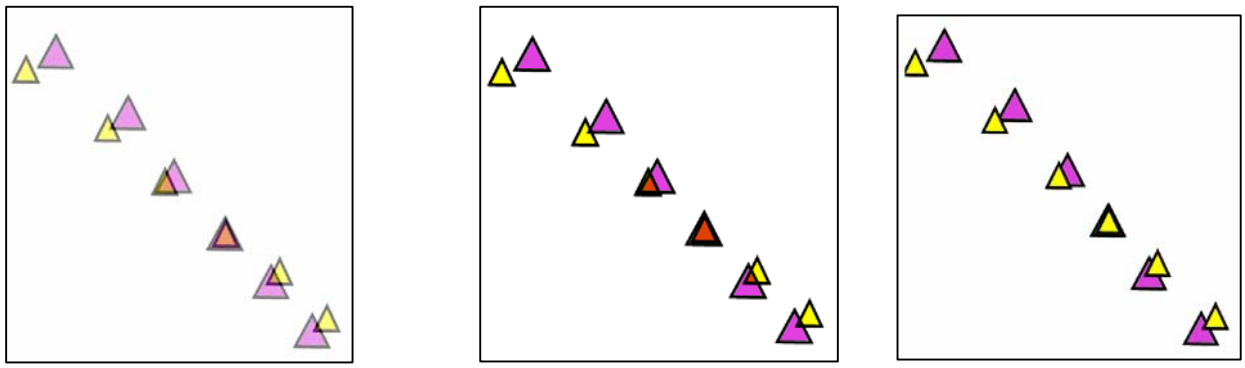


Figure 6-22: The Effect of Additive (left) and Subtractive (middle) Transparency: the Colours Combine, Easily Causing Confusion.

In this example yellow and purple combine to orange and red respectively. The picture to the RIGHT shows occlusion, the yellow triangles positioned in front occluding the purple triangles.

6.5.3.2 Actual or Potential Application to UMVs

De Vries and Jansen [33] developed a simulation environment in which human factors principles for UAV camera control are demonstrated and in which experimental studies are conducted. In an experiment they used this simulator environment to investigate the benefits of a 3D map with regard to operator performance and mental workload (Figure 6-23). They constructed a 2D map (oriented north-up) in which feedback about the control input was provided by means of an additional footprint that showed the predicted viewpoint of the camera. Operators were requested to find targets on roads and along wood edges. They could use the map to see to which part of the environment the camera was oriented. In one half of the conditions a 3D map was available together with the 2D map. The 3D map showed identical information, but was presented from the viewpoint of the camera. In some conditions, the quality of the camera images was manipulated by introducing a time delay of 1 second, or by lowering the update rate of the camera images to 3 Hz. This had no effect on the 2D and 3D map. Furthermore, in one half of the conditions a secondary task had to be performed. This was done to see whether operators had more spare mental capacity in case the 3D map was available. Several performance measures were taken. Workload was assessed using subjective and physiological measures.

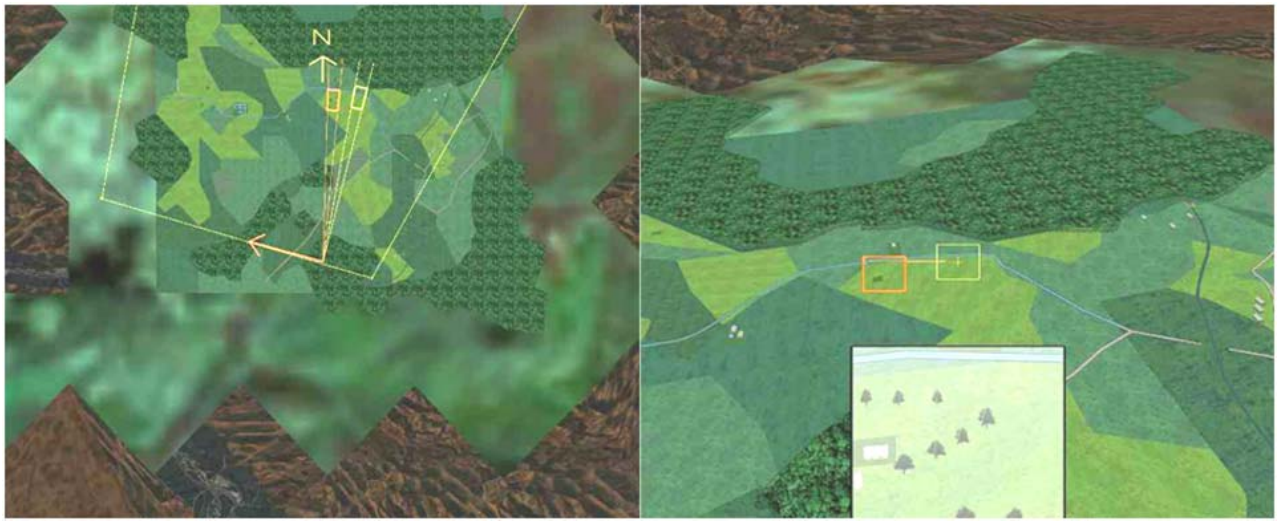


Figure 6-23: The Left Panel Shows the 2D Map with the Position of the UAV in the Center.

The right panel shows the 3D map, which is drawn from the viewpoint of the camera. In both displays, the yellow footprint shows the part of the environment that corresponds to the camera image and the orange footprint shows the predicted position based on operator input signals. The camera image is presented at the bottom of the right display.

The left panel of the simulator displayed a detailed 2D map of the environment, north-up oriented. Apart from terrain information (roads, woods, buildings, etc.) the following relevant information was presented: waypoints and the route of the UAV were shown (yellow line), flight direction of the UAV (orange arrow) and the actual and predicted footprints (see 'Footprint' below for an explanation). The right panel of the simulator displayed the 3D map, a virtual 3D world presented from the viewpoint of the UAV camera. The viewpoint for generating the 3D map depended on the camera control input of the operator, and the angular motion of this viewpoint was identical to the angular motion of the camera.

Both the 2D and 3D map showed a yellow footprint, representing the section of the map that corresponded to the camera images. The orange footprint provided direct feedback about the control input. In the conditions with low update rates and time delays, the yellow footprint followed the orange footprint. The size of the footprint could be adjusted by the zoom function, providing feedback about the zoom setting. With a zoom factor of 1, the size of the footprint on the right display was identical to the size of the camera panel. In conditions in which the 3D maps were not drawn on the right display, the footprints remained visible to provide feedback about the zoom settings.

The most important results of the experiment are presented here. The 3D map significantly improved task performance (Figure 6-24). A larger area was inspected, performance on secondary task improved, indicating that the participants had more spare mental capacity, and the participants reported lower effort.

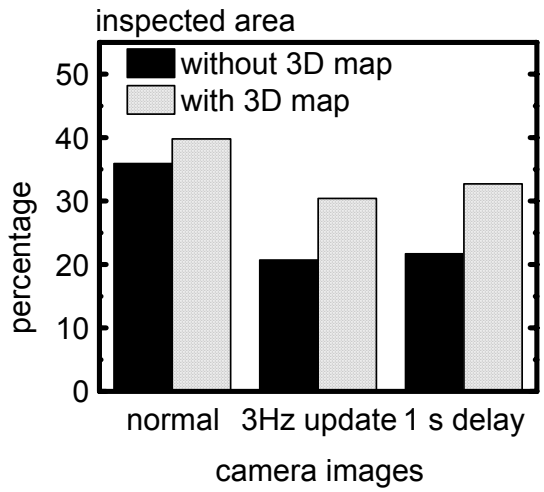


Figure 6-24: Percentage of Inspected Areas as a Function of 3D Map and the Quality of Camera Images.

The positive effect of the 3D map was largest when the quality of the camera images was low. Note that adequate feedback about time delays and low update rates was always available in both the 2D and 3D map display. Without this information, performance would be much worse in the conditions with low update rates and long time delay.

The subjective effort measure showed substantial effects as a function of all experimental factors (Figure 6-25), however, only heart rate showed a small effect as a function of the secondary task. This may be due to the lower sensitivity of physiological measures for mental effort. We found such discrepancies between subjective (Rating scale Subjective Mental Effort, RSME; [34]) and physiological measures more often [35] depending on the type of task that is evaluated. Participants most often give higher effort ratings when a task becomes more difficult as a result of a reduced quality of information.

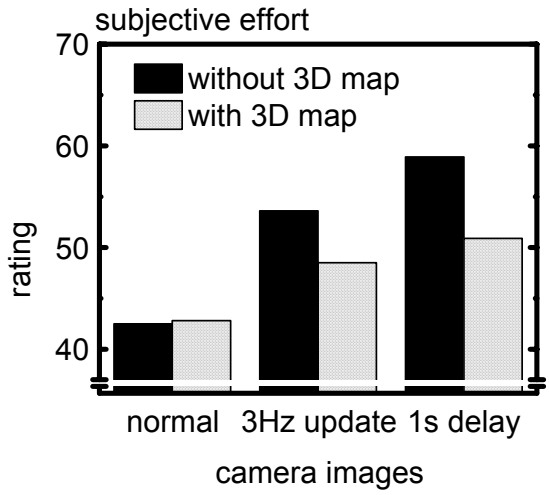


Figure 6-25: Subjective Effort Rating as a Function of Quality of the Camera Images and 3D Map.

Physiological measures most often do not show differences in these situations, because investing more effort most often will not result in better task performance. When an additional task has to be performed, attention has to be divided between more tasks. In these situations, additional effort has to be invested in order to keep an adequate level of performance of the main task. This is reflected in both subjective and physiological measures. The difference between the subjective and physiological measures can be explained along this line of thought. Degrading the quality of the display makes the task more difficult, resulting in higher effort ratings, but does not affect physiological measures because investing more effort will not improve performance. For the secondary task, more effort was required in order to maintain an adequate level of performance on the main task.

6.5.3.3 Technology Maturity, Challenges, and Unresolved Issues

3D perspective displays are used in a variety of applications. Current computing and graphic processing power allow applicability in near future UAV applications.

Although stereoscopic displays already exist in several forms, their application is still limited. The two primary bottlenecks are the associated lack of viewing comfort and the interference with other tasks. We argue that the construction of "transparent depth" displays is an important new development because it does not suffer from either of these two drawbacks. We believe the time is near to introduce transparent depth displays in professional environments like the cockpit, command and control workstations, vehicles, and hand-held devices.

6.5.3.4 References for 3D Visual Displays

[1] Wickens, C.D. and Andre, A.D. (1990). Proximity compatibility and information display: Effects of color, space, and objectness on information integration. Human Factors, 32, 61-77.

[2] Haskell, I.D. and Wickens, C.D. (1993). Two- and three-dimensional displays for aviation: A theoretical and empirical comparison. International Journal of Aviation Psychology, 3, 87-109.

[3] Wickens, C.D. and Carswell, M. (1995). The proximity compatibility principle: Its psychological foundation and relevance in display design. Human Factors, 37, 473-494.

[4] Wickens, C.D. and Prevett, T.T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays. Journal of Experimental Psychology: Applied, 1, 110-135.

[5] Wickens, C.D., Liang, C., Prevett, T. and Olmos, O. (1994). Egocentric and exocentric displays for terminal area navigation. Savoy, IL: Aviation Research Laboratory.

[6] Busquets, A.M., Parrish, R.V., Williams, S.P. and Nold, D.E. (1994). Comparison of pilot's acceptance and spatial awareness when using EFIS vs., pictorial display formats for complex, curved landing approaches. In: R.D. Gilson, D.J. Garland, and J.M. Coone (Eds.), Situational awareness in complex systems (pp. 139-170). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

[7] Grunwald, A.J., Robertson, J.B. and Hatfield, J.J. (1980). Evaluation of a computer-generated perspective tunnel display for flight-path following (Tech. Rep. No. NASA TP1736). Langley, VA: National Aeronautics and Space Administration.

[8] Roscoe, S.N. and Williges, R.C. (1975). Motion relationships in aircraft attitude and guidance displays: A flight experiment. Human Factors, 17, 374-387.

[9] Theunissen, E. (1994). Factors influencing the design of perspective flight path displays for guidance and navigation. Displays, 15, 241-254.

[10] Wickens, C.D., Liang, C., Prevett, T. and Olmos, O. (1994). Egocentric and exocentric displays for terminal area navigation. Savoy, IL: Aviation Research Laboratory.

[11] Aretz, A.J. and Wickens, C.D. (1992). The mental rotation of map displays. Human Performance, 5, 303-328.

[12] Ellis, S.R. and McGreevy, M.W. (1987). Perspective traffic display format and airline pilot traffic avoidance. Human Factors, 29, 371-382.

[13] Ellis, S.R., Kim, W.S., Tyler, M., McGreevy, M. and Stark, L. (1985). Visual enhancements for perspective displays: Perspective parameters. Proceedings of the International Conference on Systems, Man and Cybernetics (pp. 815-818). New York: The Institute of Electrical and Electronics Engineers.

[14] Barfield, W. and Young, K. (1991). Effect of geometric parameters of perspective on judgments of spatial information. Perceptual and Motor Skills, 73, 619-623.

[15] Wickens, C.D. (1992). Engineering psychology and human performance. New York: Harper Collins.

[16] McGreevy, M.W. and Ellis, S.R. (1986). The effect of perspective geometry on judged direction in spatial information instruments. Human Factors, 28, 439-456.

[17] Roscoe, S.N. and Williges, R.C. (1975). Motion relationships in aircraft attitude and guidance displays: A flight experiment. Human Factors, 17, 374-387.

[18] van Breda, L. and Veltman, J.A. (1998). Perspective information in the cockpit as a target tracking acquisition aid. Journal of Experimental Psychology: Applied, 4, 55-68.

[19] Stokes, A., Wickens, C.D. and Kite, K. (1990). Display technology: Human factors concepts. Warrendale, PA: Society of Automotive Engineers.

[20] Andre, A.D., Wickens, C.D., Moorman, L. and Boschelli, M.M. (1991). Display formatting techniques for improving situation awareness in the aircraft cockpit. International Journal of Aviation Psychology, 1, 205-218.

[21] Poulton, E.C. (1974). Tracking skill and manual control. New York: Academic Press.

[22] Dennehy, M.T., Nesbitt, D.W. and Sumey, R.A. (1994). Real-time three-dimensional graphics display for antiair warfare command and control. Johns Hopkins APL Technical Digest, Volume 15, Number 2.

[23] van Orden, K.F. and Broyles, J.W. (2000). "Visuospatial task performance as a function of two- and three-dimensional display presentation techniques". Displays, 21(1), 17-24.

[24] Pastoor, S. (1993). Human factors of 3D displays, Displays, 14.

[25] Philips. (2003). Philips 9 view auto-stereoscopic display, Website: http://www.research.philips.com/generalinfo/special/3dlcd/index.htm

[26] SeeReal Technologies d4d. Website: http://www.dresden3d.com (2003).

[27] Kooi, F.L. and Lucassen, M.P. (2001). Visual comfort of binocular and 3-D displays, In: B.E. Rogowitz and T.N. Pappas (Ed.), Human Vision and Electronic Imaging VI, pp. 586-592, The International Society for Optical Engineering, Bellingham, WA.

[28] Eagle, R., Paige, E., Rogers, B. and Sucharove, L. (2000). Accommodation cues reduce latencies for large disparity detection. Poster presented at the European Conference on Visual Perception.

[29] Soltan, P., Lasher, M., Dahlke, W., McDonald, M. and Acantilado, N. (1998). Improved second-generation 3-D volumetric display system, (Report SPAWAR Technical Report 1763), Space and Naval Warfare Systems Center, San Diego, CA.

[30] Deep Video Imaging Ltd. (2003). Interactive dual plane imagery, Website: http://www.Deepvideo.com

[31] 3dMedia. (2003). Multi-planer 3D volumetric display, Website: http://www.3dmedia.com

[32] Schoumans, N., Kooi, F.L. and Hogervorst, M.A. (2003). Binocular fusion of transparent images, (Report in preparation), TNO Human Factors, Soesterberg, The Netherlands.

[33] de Vries, S.C. and Jansen, C. (2002). Situational awareness of UAV operators onboard of moving platforms. Proceedings of the International Conference on Human-Computer Interaction in Aeronautics, pp. 144-147, HCI-Aero 2002 Cambridge, MA, 23-25 October 2002 (S. Chatty, J. Hansman, and G. Boy, Eds). Menlo Park, CA: AAAI Press.

[34] Zijlstra, F.R.H. (1993). Efficiency in Work Behavior: a design approach for modern tools. Doctoral dissertation. Delft, The Netherlands: Delft University of Technology.

[35] Veltman, J.A. and Gaillard, A.W.K. (1993). Physiological indices of workload in a simulated flight task. Biological Psychology, 42, 323-342.

6.5.4 Spatial Audio Displays

6.5.4.1 Description of Technology

Providing the appropriate information to support situation awareness is a primary challenge in the development of displays and controls for operators in any complex environment, but may be particularly challenging for the designer of interfaces for UMVs. A pilot in a traditional manned aircraft can directly perceive elements in the real world and may rapidly develop an understanding of the problem space by gleaning ambient information from peripheral factors including weather, terrain, and other vehicles in the airspace; he or she can maintain some level of understanding of the general vehicle status from displays in the cockpit, the auditory environment, and other crew members. However, many of these real-world cues are not as readily available to the operator of a UMV. Thus, the use of effective display technologies is critical to mission success.

Currently, UAV operator interfaces emphasize the presentation of information through visual displays. While often the most appropriate means for information conveyance, such systems run the risk of overloading the visual information processing capacity of the operator. The integration of display technologies focusing on other modalities affords the potential to reduce visual workload and display redundant cueing to provide for safeguards against undetected or unrecognized operationally meaningful information while also allowing for

synergistic relations to occur for the conveyance of higher-order information. Auditory display technologies, in particular, have shown great promise both as an information-bearing channel in isolation and as a component of an overall multi-modal display system. Audition excels as an early warning system – sound is inherently interpreted with respect to its signalling or warning significance. In addition, neural transmission in the auditory pathway is superior to that in the visual system, making it ideal for the display of time-critical warnings. The auditory system also plays a fundamental role in verbal communication, which is in many cases the most direct, efficient, and unambiguous means of information transfer. What is inherently appealing about these characteristics of the auditory system is that they are attention-demanding and serve to make the individual immediately aware of relevant elements in the situation. Moreover, the auditory system provides this information independent of the location of the event, for the auditory system monitors the environment from all directions at all times. Thus, information can be obtained about events in the environment even when they occur outside of the operator's visual field-of-view.

Although auditory displays do exist in most operational interfaces, they are rudimentary at best, and fail to leverage the natural spatial auditory processing capabilities of humans. That is, the ability of humans to determine the location of a sound source, monitor events at multiple locations simultaneously, and utilize auditory space as a means of organizing information, has not been fully exploited. Spatial auditory display technologies take advantage of the properties of the binaural auditory system by recreating and presenting to an operator the spatial information that would naturally be available in a "real-world" listening environment. Such displays are intuitive; they require little training and impose few additional demands on the information processing capacity of the operator.

The basic approach to generating spatialized (virtual) auditory displays assumes a "principle of equivalence" – that is, given that the display can generate a sound field at the eardrum that is identical to the sound field that would result from a real source, the perception should be equivalent to that for a real source. Thus, a virtual auditory display will result in the perception of sounds originating from real locations in space external to the listener's head.

The generation of virtual auditory displays is possible because of what is understood about the underlying mechanisms of spatial hearing and the cues that mediate sound localization. These cues arise as a result of the direction-specific, frequency-dependent modifications that are imposed on an incident waveform by a listener's head, torso, and pinnae as the sound travels from a source to a listener's eardrums. Two of these cues, interaural time differences (ITDs) and interaural level differences (ILDs), mediate sound localization in the left/right dimension. ITDs result from the fact that, due to path length differences, an acoustic waveform will arrive at the ipsilateral ear (i.e., ear nearest the sound source) earlier than at the contralateral ear (i.e., the ear on the side of the head opposite the sound source), leading to an interaural difference in time of arrival. Although these resulting interaural differences are small (< 800 μs), they are large relative to the temporal sensitivity of the auditory system, which can detect ITDs on the order of 10 μs [1]. Because of neural processing constraints, however, ITDs are primarily useful only in the low-frequency region (i.e., below about 1.5 kHz). For frequencies above approximately 3.0 kHz (a region in which the wavelength of an acoustic stimulus is small relative to the size a listener's head), the head casts an acoustic shadow such that the stimulus level at the contralateral ear is attenuated relative to the level at the ipsilateral ear. The resulting ILDs are useful for localizing sounds in this frequency region [2].

The interaural differences described above provide robust cues for sound localization in the left/right dimension; however, they are not sufficient to account for localization in elevation or front/back discrimination. Indeed, there are sets of locations, known as cones of confusion, which all produce the same interaural difference cues. A common result is that sounds will be localized to the wrong position

(e.g., the wrong elevation or front/back position) on the correct cone of confusion. The shape of the stimulus spectrum that reaches a listener's ears can help a listener disambiguate the location of a sound falling along a cone of confusion by providing cues regarding the elevation and front/back position of the sound. Narrowband spectral peaks and notches emerge from interactions between the stimulus and the pinnae in the frequency region above approximately 5.0 kHz (where the wavelength of the stimulus is small relative to the size of the pinna structures). These peaks and notches vary in relatively systematic ways as a function of elevation of the sound source, thus providing cues for localization in this dimension [3,4,5]. Additionally, each pinna casts an acoustic shadow such that a sound originating from a listener's rear hemifield will have relatively less energy in the 3.0 kHz to 6.0 kHz region than a sound originating in the frontal hemifield [6]. This filtering is presumed to help a listener resolve front/back confusions [7,8].

One additional cue that contributes to a listener's ability to determine the front/back location of a sound source arises from the fact that interaural difference cues change in predictable ways with head movement. For example, given a sound in the frontal hemifield, a clockwise rotation of the head will lead to ITDs and ILDs that indicate a sound moving to the left relative to the listener; a counter clockwise rotation in response to this same source will lead to the opposite experience. Listeners use these exploratory head movements to disambiguate the front/back location of a sound source.

All of these transformations that a waveform undergoes as it travels from a source to a listener's eardrums can be measured and captured in what is known as the head-related transfer function (HRTF; [9]). HRTFs can be measured by placing microphones in the left and right ear canals of a listener (or a mannequin head) and making recordings of broadband acoustic stimuli presented from a number of directions. Digital filters are constructed from these HRTFs and a virtual auditory stimulus can then be generated by convolving an arbitrary signal with the HRTFs for the left and right ear associated with a specific position in space. The result, when played back over headphones, is an experience in which the listener perceives the sound to have originated from that particular position in space. When done correctly, such a display can support sound localization performance that is equivalent to the localization of real sources in the free field [10]. Note that while virtual auditory displays can be delivered via loudspeaker systems instead of headphones, the veridicality of such displays is constrained by the size and acoustic characteristics of the room in which the display is presented. In addition, there is a relatively small region of space within a loudspeaker array over which a virtual auditory display will be valid for a listener. Thus, it is generally believed that headphone delivery of such displays is best. Such a practice allows the display designer to have complete control over the signal arriving at the listener's ears, independent of the conditions in which the display is presented.

In many situations, the spatial auditory display is required to present sounds that remain fixed in virtual space, independent of a listener's head movement. To accomplish this, the filters used to generate the virtual stimulus must be updated (in real-time) in response to the listener's head movement. This is achieved by continuously measuring the position of the listener's head using a head tracking system and providing this position information to the system responsible for rendering the spatial auditory stimuli. This system then convolves a sound with the filter associated with the appropriate location relative to the orientation of the listener at each point in time. Furthermore, because the HRTFs define a constrained set of discrete locations from which a virtual stimulus can be presented, an interactive display that includes head movement must be able to interpolate between sets of spatially adjacent HRTFs in order to present virtual stimuli from any arbitrary location.

A spatial auditory display with head tracking has a number of advantages. First, as discussed earlier, head movements can help a listener disambiguate the location of a sound source in the front/back dimension, thus improving localization performance. Similarly, such a display allows the listener to orient toward a

specific sound, bringing it into the "auditory fovea," where spatial acuity is greatest. However, one of the greatest advantages of an interactive display is that it can be substantially more immersive than a static display because the auditory environment reacts in realistic and expected ways in response to a listener's head movements. An interactive display is, in fact, necessary if one wishes to use virtual auditory stimuli to indicate the location of objects (such as targets and threats, as described below) in real space.

6.5.4.2 Actual or Potential Application to UMVs

As previously stated, the development of displays to support situation awareness must be a primary goal of the designer for UMV operator interfaces. Situation awareness (SA) has been defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" [11]. From this, it has been suggested that SA may be more formally broken down into three component levels that can be independently addressed, studied, and measured. They are: Level 1 SA, which refers to the perception of elements in an environment within a particular volume of time and space; Level 2 SA, which pertains to the comprehension of the meaning of these elements; and Level 3 SA, which is concerned with the projection of the status of the elements in the near future.

If one considers the types of SA-related errors that occur in aviation mishaps within the context of the UMV operator environment, it is clear that spatial audio may offer significant utility at each level of SA. The SA Error Taxonomy [12] describes several reasons that situation awareness may break down within operational settings. Level 1 SA errors have been shown to account for 76% of aviation mishaps that are attributed to human error [13]. These errors could occur as a result of the lack of availability of required data or a failure of the system to present available data. Errors might also occur because the data are provided to the operator, but are difficult to detect or are merely not observed, not attended to, misperceived, or forgotten by the operator. Spatial auditory displays would likely benefit the operator in each of these situations.

The auditory system's ability to serve as an early warning system, as well as its unique capacity for monitoring all locations simultaneously, can be exploited in spatial auditory displays. The auditory system is exquisitely sensitive to change, even when this change occurs outside of the current focus of attention. Changes associated with onsets (e.g., the introduction of new elements into a display) and offsets (e.g., the removal of an existing element from a display) are particularly well-detected. They are, in fact, often impossible to ignore, and drive the allocation of attention to elements in the environment that may yield critical information, thus reducing the chance that this information will go undetected. So natural is this phenomenon that several authors have suggested that the spatial auditory system evolved specifically to direct the visual system to meaningful events in the environment. Support for this hypothesis can be seen in laboratory studies involving visual search tasks, where spatially coincident auditory cues have been shown to reduce visual target acquisition and identification times by a factor of 2-5 in very simple visual scenes; much greater benefits have been found in more complex visual scenes such as those depicted in simulated operational environments [14,15,16]. Moreover, redundantly coding information using both auditory and visual stimuli may help to overcome systematic misperceptions. That is, the display of auditory information that is consistent with, and covaries with, visual information is not only unambiguous, but is consistent with operator expectancies, thus providing a more natural, intuitive interface. Given these issues, it seems clear that spatial audio cueing might be especially useful to UMV operators who are tasked with finding ground targets in a remote environment through the control of a maneuverable camera.

The success of future UMV systems will also depend on an operator's ability to monitor multiple independent sources of information in order to maintain situation awareness. Here, again, the auditory modality is particularly well suited to the task, for the auditory system is capable of segregating multiple simultaneous

sounds into different streams [17], thus allowing an observer to attend to the most relevant information at any given time and to relegate the remaining information to the "background" for further processing when necessary. This ability to focus on certain information, while simultaneously and in parallel monitoring and maintaining a functional level of awareness about other sources of information, is invaluable in complex and dynamic operational environments. Because this segregation process is, in part, mediated by space [18], it can be exploited using spatial auditory displays.

One environment in which the utility of spatial audio for segregation has been extensively studied is that of the operator who must monitor multiple simultaneous communication channels. Here, substantial improvements in workload and communication effectiveness can be seen when the individual communication channels are presented such that they appear to originate from different locations in space, unlike standard monaural communication systems [19,20,21]. The flexibility of this display technology allows a user to configure the display such that particular communication channels are assigned to specific locations that are meaningful to the user. The UMV operator who must engage in verbal communications with a variety of distributed team members could make use of such a display to maintain an awareness of the identity of the specific talker from which each communication originated. That is, the operator can use space as an organizing principle, and this may result in enhancements of both Level 1 and Level 2 SA.

The appropriate development and comprehension of mental models necessary for achieving Level 2 SA may also be supported by an auditory environment in which the operator is immersed and experiences a sense of presence (i.e., "being there"). Support for this comes from the work of Ramsdell, who suggested that one function of the auditory system is to connect one to the real world on a primitive level, utilizing the incidental sounds that serve to make up the auditory "background." [22]. This work was based on reports from suddenly-deafened individuals who reported that the world seemed "dead" and "(un)coupled," that "it was almost impossible to believe in the passage of time …couldn't hear a clock tick" (p.503). Ramsdell distinguished this level of auditory perception from that of communication and warning, which are more obvious and overt functions of the auditory system, and suggested that this primitive level of perception is critical for a sense of "connectedness" to the world. This experience of suddenly-deafened individuals has been likened to that of the user of a virtual environment with an impoverished auditory display [23]. Perceptually rich virtual auditory environments are believed to lead to a strong sense of presence [24]. Although the link between presence and task performance is less clear [25] than that believed to exist between situation awareness and performance, it has been suggested that this is due in part to the lack of a robust measure of presence and/or the use of gross performance metrics that may not be sensitive to issues regarding how the interface is actually being used [26], and thus how the sense of presence may contribute to that usage. The sense of presence is concomitant with an engagement on the part of the operator, and this may be critical when the operator takes on a supervisory role over semi-autonomous UMVs. In this situation, there exists the potential that the operator will 'fall out' of the control loop and may have difficulty re-entering when necessary. Immersion in the virtual environment (i.e., the UMV operator interface) may facilitate intuitive interaction and ensure that the operator remains engaged in the mission even if not directly flying the vehicle.

Finally, the support of Level 3 SA may be assisted by an auditory display that is spatially, spectrally, and temporally dynamic. Information about the current and future states of highly-dimensional environments may be related via auditory information in a way that is engaging and intuitive. Operators may discern overall relationships and trends in order to better predict future states [27]. Auditory motion perception can be used to demonstrate trajectories of elements in the environment and are particularly compelling when used in conjunction with analogous visual displays for predicting future states, allowing the UMV operator to "fly several seconds ahead of the aircraft." This temporal aspect of situation awareness may be particularly well-supported by a spatial auditory display.

6.5.4.3 Technology Maturity, Challenges, and Unresolved Issues

Many of the technological challenges that were considered in early iterations of spatial auditory displays have been largely overcome, the most obvious of which being adequate computational power. Increased memory capacity has allowed for the storage of a large number of measured HRTF locations, thus increasing the resolution of the rendered auditory space. Advances in digital signal processing capabilities have allowed for the rapid convolution of stimuli with more detailed representations of the HRTFs, and thus the presentation of increasingly complex auditory environments (e.g., multiple sources, room acoustics characteristics) that have the potential to yield richer and more compelling experiences for UMV operators.

Perhaps the greatest challenges that remain in the generation of veridical and compelling spatial auditory displays concern the psychoacoustic questions that continue to be of interest to auditory scientists. One issue concerns the fact that HRTFs are highly individualized, and it has been shown that the localization of virtual auditory stimuli is best when listening through one's own HRTFs. This is especially true for localization in those dimensions in which performance is mediated by spectral cues (i.e., elevation and front/back dimensions). However, because it is, in many cases, logistically impossible to collect HRTF measurements on each individual operator, the most common practice involves the use of generic HRTFs that were originally measured using a mannequin head. A better understanding of the relative importance of various spectral features for localization, and how these features are recovered by the auditory system, might allow for the construction of an effective set of non-individualized HRTFs. Such HRTFs could improve performance over existing generic models of auditory space and perhaps yield localization performance comparable to that found for free-field stimuli (i.e., greater accuracy, fewer front/back confusions, greater externalization of the auditory images).

Another issue that must be addressed in future developments of spatial auditory displays is the encoding of veridical sound source distance. For an arbitrary stimulus, there exist several acoustic cues that yield source distance information, among them sound source intensity, gross spectral characteristics, and the ratio of direct-to-reverberant acoustic energy reaching a listener [28,29]. It has been suggested that listeners utilize these cues when determining the distance of a remote sound source, thus dictating that such cues be incorporated into displays lest one provide impoverished distance information. In addition, in the auditory near field (i.e., less than 1 m from the head), distance perception is mediated in large part by characteristics in the HRTF that vary with source distance, in particular interaural level differences [30]. These cues appear to provide relatively salient cues about sound source distance, and thus should be implemented in future displays. However, auditory displays are likely to be used in noisy environments, and noise serves to disrupt localization cues not only in azimuth and elevation [31], but can also disrupt the cues for source distance by masking the reverberation and spectral cues in the display, and by reducing the dynamic range utilized for intensity-based distance coding. New symbologies must be developed that could be used in conjunction with existing cues to provide more reliable sound source distance. For example, one display that has been proposed [32] utilizes the effort with which a speech phrase is spoken to indicate distance – shouted speech indicates sources at greater distances, whispered speech indicates closer sources, and conversational speech indicates sources at distances in between.

It is important to note that spatial auditory displays have been integrated and tested in a number of operational environments including fighter jets and general aviation aircraft, as well as command and control centers [33,34]. The Air Force Research Laboratory has implemented a spatial audio display in two ground-based controller stations as part of a communications system upgrade at a training facility. This system, which allows users to allocate incoming voice communications to seven different apparent locations, has made it possible for dozens of operators to experience the advantages of spatial audio during the conduct of realistic

training missions. Operator feedback indicates that these spatial auditory displays allowed the operators to have "total SA" in the mission. Their comments indicate not only overwhelming acceptance by the user community, but also the increase in mission effectiveness that is possible due to the enhanced situation awareness supported by spatial audio. Spatial audio can therefore be classified as having a Technology Readiness Level of 8.

There are a number of unexplored applications for spatial audio that have the potential to enhance situation awareness for UMV operators. The need to monitor multiple simultaneous environments (e.g., the virtual operational environment and the real-world environment in which the operator station is located) may be supported by signal processing techniques employing room acoustics models to make the two categories of display elements appear to originate from different "rooms." An auditory environment that is slaved to the UMV camera may allow the operator to unambiguously center a visual target in a complex visual scene that would otherwise be difficult to find. Spatial audio displays can also lead to a level of realism that as yet cannot be achieved in visual displays. Thus, they contribute substantially to a sense of presence and task engagement that could potentially improve overall operator performance. The challenge is to identify those specific features that contribute to presence and implement them. Nevertheless, in the nearly 20 years since the first functional spatial auditory display system was introduced, great advances have been made in both science and technology, resulting in a display that is reliable, mature, and cost-effective.

6.5.4.4 References for Spatial Audio Displays

[1] Klump, R.G. and Eady, H.R. (1956). Some measurements of interaural time difference thresholds. Journal of the Acoustical Society of America, 28, 859-860.

[2] Feddersen, W.E., Sandel, T.T., Teas, D.C. and Jeffress, L.A. (1957). Localization of high-frequency tones. Journal of the Acoustical Society of America, 29, 988-991.

[3] Blauert, J. (1969/1970). Sound localization in the median plane. Acustica, 22, 205-213.

[4] Butler, R.A. and Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. Journal of the Acoustical Society of America, 61, 1264-1269.

[5] Middlebrooks, J.C. (1992). Narrow-band sound localization related to external ear acoustics. Journal of the Acoustical Society of America, 92, 2607-2624.

[6] Shaw, E.A.G. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. Journal of the Acoustical Society of America, 56, 1848-1861.

[7] Gardner, M.B. and Gardner, R.S. (1973). Problem of localization in the median plane: Effect of pinnae cavity occlusion. Journal of the Acoustical Society of America, 53, 400-408.

[8] Oldfield, S.R. and Parker, S.P.A. (1984). Acuity of sound localization: A topography of auditory space. I. Normal hearing conditions. Perception, 13, 581-600.

[9] Wightman, F.L. and Kistler, D.J. (1989a). Headphone simulation of free-field listening. I: stimulus synthesis. Journal of the Acoustical Society of America, 85, 858-867.

[10] Martin, R.M., McAnally, K.I. and Senova, M.A. (2001). Free field equivalent localization of virtual audio. Journal of the Audio Engineering Society, 49, 14-22.

[11] Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 97-101). Santa Monica, CA: Human Factors Society.

[12] Endsley, M.R. (1995). A taxonomy of situation awareness errors. In: R. Fuller, N. Johnston, and N. McDonald (Eds.), Human factors in aviation (pp. 287-292). Aldershot, England:Ashgate Publishing Ltd.

[13] Jones, D.G. and Endsley, M.R. (1996). Sources of situation awareness errors in aviation. Aviation, Space, and Environmental Medicine, 67, 507-512.

[14] Bolia, R.S., D'Angelo, W.R. and McKinley, R.L. (1999). Aurally-aided visual search in three-dimensional space. Human Factors, 41, 664-669.

[15] Simpson, B.D., Bolia, R.S., McKinley, R.L. and Brungart, D.S. (2005). The impact of hearing protection on sound localization and orienting behavior. Human Factors, 47(1), 188-198.

[16] Bronkhorst, A.W., Veltman, J.A. and van Breda, L. (1996). Application of a Three-Dimensional Auditory Display in a Flight Task. Human Factors, 38, 23-33.

[17] Bregman, A.S. (1990). Auditory scene analysis. Cambridge, MA: MIT Press.

[18] Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and two ears. Journal of the Acoustical Society of America, 25, 975-979.

[19] Drullman, R. and Bronkhorst, A.W. (2000). Multichannel speech intelligibility and speaker recognition using monaural, binaural, and 3D auditory presentation. Journal of the Acoustical Society of America, 107, 2224-2235.

[20] Brungart, D.S., Ericson, M.A. and Simpson, B.D. (2002). Design considerations for improving the effectiveness of multitalker speech displays. Proceedings of ICAD 2002 (pp. 424-430). International Community for Auditory Display.

[21] Bolia, R.S., Nelson, W.T., Vidulich, M.A., Simpson, B.D. and Brungart, D.S. (2005). Communications research for command and control: Human-machine interface technologies supporting effective air battle management. Proceedings of the 10th International Command and Control Research and Technology Symposium. Washington: Command and Control Research Program.

[22] Ramsdell, R.S. (1978). The psychology of the hard-of-hearing and the deafened adult. In: H. Davis and S.R. Silverman (Eds.), Hearing and deafness (4th ed., pp. 499-510). New York: Holt, Rinehart & Winston.

[23] Gilkey, R.H. and Weisenberger, J.M. (1995). The sense of presence for the suddenly deafened adult. Presence, 4, 357-363.

[24] Gilkey, R.H., Simpson, B.D. and Weisenberger, J.M. (2001). Creating Auditory Presence. Proceedings of the HCI 2001 International Conference, New Orleans, 609-613.

[25] Welch, R.B. (2000). How can we determine if the sense of presence affects task performance? Presence: Teleoperators and Virtual Environments, 9, 574-577.

[26] Kalawsky, R.S. (2000). The validity of presence as a reliable human performance metric in immersive environments. Proceedings of the 3rd International Workshop on Presence, Delft, The Netherlands.

[27] Kramer, G. (1994). An introduction to auditory display. In: G. Kramer (Ed.), Auditory display: Sonification, audification, and auditory interfaces (pp. 1-77). Reading, MA: Addison-Wesley.

[28] Zahorik, P. (2002). Auditory display of sound source distance. Proceedings of the 2002 International Conference on Auditory Display.

[29] Bronkhorst, A.W. and Houtgast, T. (1999). Auditory distance perception in rooms. Nature, 397, 517-520.

[30] Brungart, D.S. and Rabinowitz, W.M. (1999). Auditory localization of nearby sources. Head-related transfer functions. Journal of the Acoustical Society of America, 106, 1465-1479.

[31] Good, M.D. and Gilkey, R.H. (1996). Sound localization in noise: The effect of signal-to-noise ratio. Journal of the Acoustical Society of America, 99, 1108-1117.

[32] Brungart, D.S. (2000). A speech-based auditory distance display. Proceedings of the 109th Convention of the Audio Engineering Society, 22-25.

[33] McKinley, R.L. and Ericson, M.A. (1997). Flight demonstration of a 3-D auditory display. In: R.H. Gilkey and T.R. Anderson (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments (pp. 683-699). New Jersey: Lawrence Erlbaum Associates.

[34] Simpson, B.D., Brungart, D.S., Dallman, R.C., Joffrion, J., Presnar, M.D. and Gilkey, R.H. (2005). Spatial audio as a navigation aid and attitude indicator. Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, 1602-1606.

6.5.5 Haptic Display Technology

6.5.5.1 Description of Technology

Designers of human machine interfaces are increasingly applying multi-modal interfaces. An important reason for this is the need for an alternative or complementary information channel in complex operator environments [1,2,3]. Traditionally, the auditory channel is often considered as an alternative or supplement to visual displays. Examples include the presentation of route navigation [4,5] and tracking error information [6,7]. However, there are situations in which the visual and auditory channels of an operator are both heavily loaded or in which the visual and/or auditory information is degraded. In those situations, a haptic display system may be useful.

An example of a tactile display is the TNO tactile waist belt (Figure 6-26). The tactile display consists of eight vibrating elements (1.3 V vibrating DC motors, housed in rectangular PVC boxes) with a body contact area of 1.5 by 2.0 cm and a vibration frequency of 155 Hz. The boxes are mounted in an adjustable waist belt. The resolution of the displays (i.e., 8 tactors for 360°) is in between the minimum required (i.e., two elements: one for left and one for right) and the limit of direction perception on the torso (to be in the order of 10°). The location of the elements in the belt is adjustable so they can easily be positioned in the direction of the cardinal and oblique axes irrespective of the body form of the person who is wearing the belt. The waist belt is worn over the underclothing of the person.

Figure 6-26: Example Tactile Display (Tactile Waist Belt Clearly Showing the Vibration Elements).

Another example is the tactile torso display. TNO [8,9] used a torso display consisting of 64 vibro-tactile elements that presented information concerning the desired direction of motion (simple version), and a torso display that included information concerning an actual motion direction (complex version). Figure 6-27 shows such a torso display used in experiments based on helicopter scenarios.

Figure 6-27: Example of a Tactile Torso Display Used in Helicopter Orientation Studies [21].

6.5.5.2 Actual or Potential Application to UMVs

6.5.5.2.1 Vibrotactile Displays

Vibrotactile displays present information by delivering a localised vibration to the skin. Partly due to the trends in multi-modal interfaces, this kind of display is going through a rapid development. Although the technology to build active displays was already developed 40 years ago, the applications were mainly restricted to research tools. An important example is the TVSS, a device developed by Bach-y-Rita and

colleagues that displays visual information acquired by a camera on a 144-element vibrotactile display attached to the abdomen of the observer [10]. Today, cellular phones with vibration function are probably the best-known example of a tactile display, but many examples of tactile displays are also developed within the military domain. In relation to UAVs, tactile displays are receiving a small share of interest, as far as we know they are only investigated in several laboratories, but are not on the market yet.

We will discern the following two areas in which adding a tactile display to the UMV-operator interface may be beneficial.

a) (Directional) warning/attention allocation system. The vibration function on a mobile phone and the stick shaker, which warns pilots that the aircraft is in danger of stalling, are two examples of a tactile warning system. Experimental evidence that indicates that vibrotactile displays are well suited to detect time critical events is provided by several authors. For example, Martens and Van Winsum [11] found that tactile warning cues were more effective than speech warning cues in presenting collision avoidance warnings. Sklar and Sarter [12] demonstrated that the use of tactile cues for indicating unexpected changes in status are more effective than visual cues. Additionally, some research has been conducted on the usefulness of vibrotactile displays within UAV control applications [13,14,15]. The UAV operator wore small tactors mounted in elastic bands, one on each inner wrist (Figure 6-28). When a high priority system contingency occurred, one or both of the tactors vibrated. By noting which tactor(s) was vibrating, the operator could rapidly identify the system that needs attention, especially when attention was directed to a display not containing a visual warning. This technology has strong potential for directing attention to unexpected events in future UMV control environments where the operator supervises automated systems.

Figure 6-28: Tactile Wrist Pads as High Priority Alert Cue.

b) Spatial information. This category is of special interest to UMV applications and also the area that receives attention for military applications. Since the nineties, the tap-on-the-shoulder principle is implemented in tactile torso displays. The power behind this concept is that the proverbial tap on shoulder draws and directs the spatial attention of the observer. Displays that can provide this tap on any location on the torso and that consider the torso as a 3D sphere are able to project 3D spatial information directly and intuitively. Several preconditions for successful application of this principle have been fulfilled, e.g., the spatial resolution and the accuracy of direction perception have been investigated [16], as are the effects of applying the principle in dynamic environments and in combination with other sensory modalities [17,18]. Proof-of-concept studies also show encouraging results for helicopter hover tasks [19,20,21], spatial disorientation situations [22,23,24], waypoint navigation [16] and vehicle control [24,25]. In UMVs, acquiring spatial information or a good sense of spatial awareness is one of the critical issues. For example, the multiple frames of reference for an operator (especially a moving operator, i.e., an operator that is on board a moving platform) may slow

down this process and the quality of the decisions made. A tactile display that uses the tap-on-the-shoulder principle can be a powerful situation awareness support [26,27,28], for instance by presenting the heading direction of the vehicle.

6.5.5.2.2 Force Feedback Displays

Force feedback (proprioceptive) displays present information via forces on the controls. Force feedback devices can be used in a general human-computer interaction setting (e.g., a force feedback mouse or rollerball), but are not particularly widespread. On the other hand, force feedback joysticks and steering wheels are quite common in gaming. Also, force displays have proven their value in many remote control situations. In UMVs, force feedback devices can be also be of use, amongst others because the control of their inhabited equivalents relies on these devices. For example, in road vehicles, forces (and vibrations) in the steering wheel provide important information about the road conditions as well as the vehicle behaviour. The same holds for forces in flight controls of helicopters and airplanes. Especially when the UMV is controlled in the loop, this information may be of great value. Force displays may also assist operators in camera control tasks, for example by presenting platform motions via joystick motions or forces [29,30].

Wind turbulence is potentially detrimental to safe and effective UAV tele-operated control. Unfortunately, the physical separation of the crew from the aircraft makes detection of sudden turbulence onset very difficult, often solely indicated by an unexpected perturbation of video images transmitted from a UAV mounted camera. One study [31] explored how to provide the UAV pilot with an enhanced indication of turbulence. Four different alerts were evaluated: Visual (perturbation of nose-camera imagery and overlaid HUD symbology – Baseline), Visual/Haptic (Visual and additional 1 second, low gain, high frequency vibration of the control stick using an Immersion Corporation 2000 Force Feedback Joystick), Visual/Aural (Visual and 1 second pure tone), Visual/Aural/Haptic (all three cues simultaneously). Data were collected from pilots as they performed simulated landing tasks. Conditions containing the haptic cue (Visual/Haptic and Visual/ Haptic/Aural) resulted in less error than non-haptic cue conditions (Visual and Visual/Aural). Although the aural alert also improved landing accuracy and detection of turbulence direction, performance was best with the redundant kinesthetic feedback. When randomly queried regarding the primary direction of the UAV immediately following a turbulence event, participants were more accurate when haptic feedback had been present. It should be noted that the operators commented that only a slight haptic feedback is required (and desired) to alert turbulence onset.

6.5.5.3 Technology Maturity, Challenges, and Unresolved Issues

The technology for haptic feedback is very mature, though experiments designed to determine the effectiveness of haptic information for military applications are on-going. Early studies show promise for the effective use of this technology, with practical military applications on the near horizon.

6.5.5.4 References for Haptic Display Technology

[1] Spence, C. and Driver, J. (1997). Cross modal links in attention between audition, vision, and touch: implications for interface design. International Journal of Cognitive Ergonomics, 1 (4), 351-373.

[2] Wickens, C.D. and Hollands, J.G. (2000). Attention, time-sharing, and workload. Chapter 11 in Engineering Psychology and Human Performance. NY: Prentice Hall.

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

[4] Parkes, A.M. and Coleman, N. (1990). Route guidance systems: A comparison of methods of presenting directional information to the driver. In: E.J. Lovesey (Ed.), Ergonomics – Setting standards for the '90's. London: Taylor & Francis, pp. 480-485.

[5] Streeter, L.A., Vitello, D. and Wonsiewicz, S. (1986). How to tell people where to go: comparing navigational aids. International Journal of Man-Machine Interaction, 22, 549-562.

[6] Forbes, T.W. (1946). Auditory signals for instrument flying. Journal of Aeronautical Science, 13, 255-258.

[7] Wickens, C.D. (1992). Engineering psychology and human performance. New York: Harper Collins.

[8] van Erp, J.B.F. (2005). Presenting Directions with a Vibro-Tactile Torso Display. Accepted by Ergonomics.

[9] van Erp, J.B.F., Groen, E.L., Bos, J.E. and Van Veen, H.A.H.C. (2005). A Tactile Cockpit Instrument Supports the Control of Self-Motion During Spatial Disorientation. Human Factors.

[10] Bach-y-Rita, P., Collins, C.C., Saunders, F., White, B. and Scadden, L. (1969). Vision substitution by tactile projection. Nature, 221.

[11] Martens, M.H. and Van Winsum, W. (2001). Effects of speech versus tactile support messages on driving behaviour and workload. Proceedings of the 17th International Technical Conference on Enhanced Safety of Vehicles, Amsterdam.

[12] Sklar, A.E. and Sarter, N.B. (1999). Good vibrations: tactile feedback in support of attention allocation and human automation coordination in event-driven domains. Human Factors, 41 (4), 543-452.

[13] Calhoun, G.L., Draper, M.H., Ruff, H.A. and Fontejon, J.V. (2002). Utility of a tactile display for cueing faults. Proceedings of the Human Factors and Ergonomic Society 46th Annual Meeting (pp. 2144-2148), Santa Monica, CA: Human Factors and Ergonomic Society.

[14] Calhoun, G.L., Draper, M.H., Ruff, H.A., Fontejon, J.V. and Guilfoos, B. (2003). Evaluation of tactile alerts for control station operation. Proceedings of the Human Factors and Ergonomic Society 47th Annual Meeting (pp. 2118-2122), Santa Monica, CA: Human Factors and Ergonomic Society.

[15] Calhoun, G.L., Ruff, H.A., Draper, M.H. and Guilfoos, B.J. (2005). Tactile and aural alerts in high auditory load UAV control environments. Proceedings of the Human Factors and Ergonomic Society 49th Annual Meeting (pp. 145-149), Santa Monica, CA: Human Factors and Ergonomic Society.

[16] van Erp, J.B.F., van Veen, H.A.H.C., Jansen, C. and Dobbins, T. (2005). Waypoint Navigation with a Vibrotactile Waist Belt. Accepted by Transactions on Applied Perception.

[17] van Erp, J.B.F. and Werkhoven, P.W. (2004). Vibro_tactile and visual asynchronies: Sensitivity and consistency. Perception, 33, 103-111.

[18] van Erp, J.B.F. and Verschoor, M.H. (2004). Cross-Modal Visual and Vibro-Tactile Tracking. Applied Ergonomics, 35, 105-112.

[19] Rupert, A.H. (2000). An instrumentation solution for reducing spatial disorientation mishaps. IEEE Engineering in Medicine and Biology, 19, 71-80.

[20] Rupert, A.H. (2000). Tactile situation awareness system: proprioception prostheses for sensory deficiencies. Aviation Space and Environmental Medicine, 71, A92-A99.

[21] van Erp, J.B.F., Veltman, J.A., Van Veen, H.A.H.C. and Oving, A.B. (2003). Tactile Torso Display as Countermeasure to Reduce Night Vision Goggles Induced Drift. Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures. RTO Meeting Proceedings, 86, pp. 49-1 – 49-8. Neuilly-sur-Seine Cedex, France: NATO RTO.

[22] Rupert, A.H., Guedry, F.E. and Reschke, M.F. (1993). The use of a tactile interface to convey position and motion perceptions. AGARD meeting proceedings on 'Virtual interfaces: research and applications'. Neuilly-sur-Seine, France: RTO NATO.

[23] Benson (2003). Technical Evaluation Report. Spatial disorientation in military vehicles: causes, consequences and cures. RTO Meeting Proceedings 86. Neuilly-sur-Seine Cedex, France: NATO RTO.

[24] van Erp, J.B.F. and Van Veen, H.A.H.C. (2004). Vibrotactile in vehicle navigation system. Transportation Research, Human Factors, 7, 247-256.

[25] Dobbins, T. and Samway, S. (2002). The use of tactile navigation cues in high-speed craft operations. Proceedings of the RINA Conference on High Speed Craft: technology and operation, pp. 13-20. London: The Royal Institution of Naval Architects.

[26] Gilliland, K. and Schlegel, R.E. (1994). Tactile stimulation of the human head for information display. Human Factors, 36 (4), 700-717.

[27] Raj, A.K., Kass, S.J. and Perry, J.F. (2000). Vibrotactile displays for improving spatial awareness. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Santa Monica, CA: The Human Factors and Ergonomics Society, pp. 181-184.

[28] Raj, A.K., McGrath, B.J., Rochlis, J., Newman, D.J. and Rupert, A.H. (1998). The application of tactile cues to enhance situation displays. 3rd Annual Symposium & Exhibition on Situational Awareness in the Tactical Air Environment, (pp. 77-84). Patuxent River, MD.

[29] Korteling, J.E. and Van der Borg, W. (1997). Partial camera automation in an unmanned air vehicle. IEEE Transactions on Man and Cybernetics A, 27 (2), 256-262.

[30] Korteling, J.E. and Van Emmerik, M.L. (1998). Continuous haptic feedback in target tracking from a moving platform. Human Factors, 40 (2), 198-208.

[31] Draper, M.H., Ruff, H.A., Repperger, D.W. and Lu, L.G. (2000). Multi-sensory interface concepts supporting turbulence detection by UAV controllers. Proceedings of the Human Performance, Situational Awareness and Automation Conference (pp. 107-112). Savannah Georgia.

6.6 INTERFACE ISSUES FOR MULTI-UMV SUPERVISORY CONTROL

6.6.1 Implications of Automation/Autonomy

6.6.1.1 What is Automation and Supervisory Control?

A common frame of reference is needed to discuss automation and subsequently supervisory control. The Oxford Dictionary of Current English [1] defines automation as "use or introduction of automatic methods or equipment in place of manual labour." Thus, automation takes place when a task that is usually performed by a human is performed by a machine (often a computer). An automatic transmission on a car, for example, takes the place of a human physically shifting the gears, and automates this task. The process by which the human controls the automated system (i.e., selecting the appropriate gear) is supervisory control [2]. This section will focus on the interaction between automation and supervisory control or more specifically, the interface implications posed by automation.

Automation is not all or none. It comes in many forms and varieties and has been characterized as levels of automation or supervisory control. Sheridan and Verplank [3] proposed one of the first automation taxonomies characterized by eight levels of supervisory control (Table 6-1).

Table 6-1: A Scale of Degrees of Automation (Sheridan and Verplank, 1978)

1	The computer offers no assistance; the human must do it all.
2	The computer suggests alternative ways to do the task.
3	The computer selects one way to do the task and executes that suggestion,
4	if the human approves, or
5	allows the human a restricted time to veto before automatic execution, or
6	executes automatically, then necessarily informs the human, or
7	executes automatically, then informs the human only if asked.
8	The computer selects the method, executes the task and ignores the human.

More recently Parasuraman, Sheridan and Wickens [4] developed a framework for making decisions on what functions should be automated and to what extent. Four classes of function were proposed:

1) Information acquisition;
2) Information analysis;
3) Decision and action selection; and
4) Action implementation.

Automation can vary within each function from low to high (see Table 6-2) and a particular system can be automated at different levels in each of the four functions. The model then presents a methodology for deciding what level of automation should be assigned what level of automation for a particular system. This methodology is based on primary criteria; mental workload, situation awareness, and complacency, as well as secondary criteria: automation reliability and costs of decisions/outcomes.

Table 6-2: Levels of Automation of Decision and Action Selection [4]

HIGH	10	The computer decides everything, acts autonomously, ignoring the human.
	9	informs the human only if it, the computer, decides to
	8	informs the human only if asked, or
	7	executes automatically, then necessarily informs the human, and
	6	allows the human a restricted time to veto before automatic execution, or
	5	executes the suggestion if the human approves, or
	4	suggests one alternative
	3	narrows the selection down to a few, or
	2	The computer offers a complete set of decision/ action alternatives, or
LOW	1	The computer offers no assistance: human must make all decisions and actions.

6.6.1.2 Levels of Automation Specific to UMVs

These general taxonomies of automation have served well for research and manufacturing systems. The unique nature of UMVs, though, has led to new characterizations of levels of automation and the relationship between the operator and the UMV. No universally agreed upon description of levels of automation has emerged, but key concepts are similar in different taxonomies. Two examples are provided below. The U.S. Office of the Secretary of Defence [5] developed a UAV roadmap and included the following levels of automation, with a proposed timeline (Figure 6-29). Note that an exponential increase in the levels of automation is predicted. In 1995, automation was at level 2, Real Time Health/Diagnosis, progressing to level 4, Onboard Route Re-planning by 2005. However, by 2015, it is predicted that automation will reach level 10, Fully Autonomous Swarms. This will require great leaps of technology and perhaps greater leaps of our understanding of how humans interact and control autonomous systems.

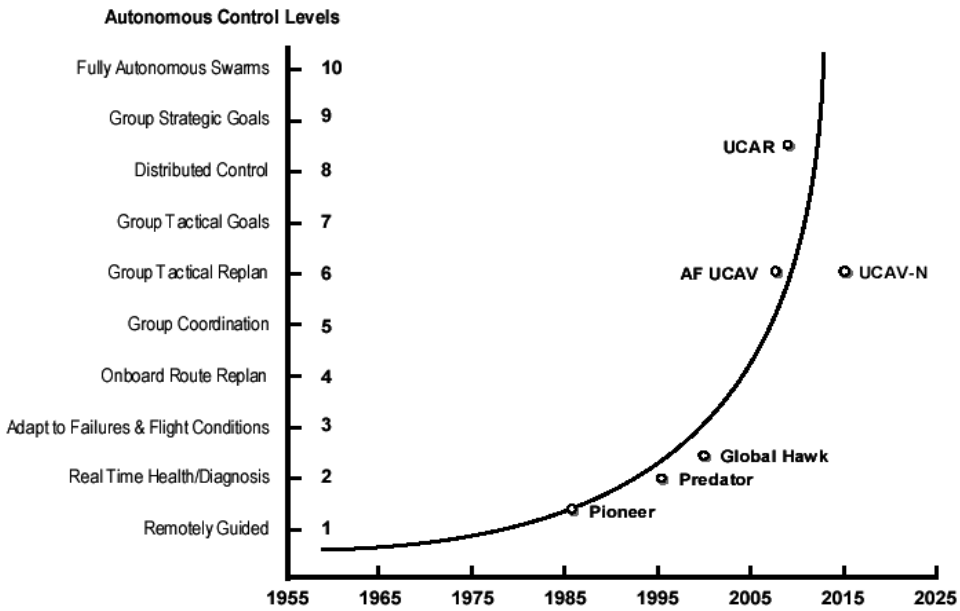


Figure 6-29: Levels of Automation [5].

Another conceptualisation of levels of automation comes from the U.S. Army Science Board [6] (Figure 6-30).

Figure 6-30: Levels of Automation, U.S. Army Science Board [6].

While less specific about time frame, the rapid increase in automation is also evident in the Army Science Board taxonomy.

In addition to the various levels of autonomy taxonomies, a distinction needs to be made concerning the levels of interoperability and control. These levels were defined in a recent NATO Standardisation Agreement or STANAG [7], as outlined in Table 6-3.

Table 6-3: Levels of Interoperability

Level I	Indirect receipt of imagery and/or data.
Level II	Direct receipt of imagery and/or data; where direct covers of reception of the UAV payload data by the UCS when it has direct line-of-sight with the UAV or a relay device which has direct line-of-sight with the UAV.
Level III	Control and monitoring of the UAV payload in addition to direct receipt of imagery/data.
Level IV	Control and monitoring of the UAV, less launch and recovery.
Level V	Control and monitoring of the UAV (Level IV), plus launch and recovery functions.

A common feature in the levels of automation taxonomies is that the level of autonomy is predicted to rise dramatically over the next several years. This rise in autonomy will dramatically affect the relationship of the operator to the aircraft. An interface designer needs to carefully examine the required or optimal levels of autonomy and the levels of control. The rest of this chapter addresses the implications of this rising automation level.

6.6.1.3 Automation: A Double-Edged Sword

6.6.1.3.1 Potential Benefits of Automation

Why do we automate? There are several benefits, both real and perceived. However, there are also costs (or potential costs) to automation and this section will address these costs and benefits.

6.6.1.3.1.1 Mission Capability

A commonly cited benefit of using UMVs is that they can do the jobs that humans are either reluctant to do or cannot do. These missions have been referred to as “dirty, dangerous, and dull”. UMVs can clean up toxic spills, or otherwise dirty environments without exposing human operators to this potential danger. UMVs might have saved lives at Chernobyl. About 200,000 people (“liquidators”) from all over the USSR were involved in the recovery and clean up of this accident during 1986 and 1987. They received high doses of radiation, around 100 millisieverts. Some 20,000 of them received about 250 mSv and a few received 500 mSv. Later, the number of liquidators swelled to over 600,000, but most of these received only low radiation doses [8]. UMVs might have saved lives by performing this “dirty” mission. UMVs can also operate in combat environments without exposing the operators to danger. Some visions of the future suggest that manned war as we know it won’t exist in the future. We will instead, see battles of drone aircraft and land vehicles. Automated vehicles are also very useful for performing the very dull, boring missions. These might include long duration surveillance or long egress segments. UAVs are making strides in development of this type of capability. Within two years, solar-electric airplanes incorporating energy storage for night-time operations will be capable of continuous flight for up to six months at a time at altitudes over 60,000 feet. Applications for such aircraft include telecommunications, remote sensing and atmospheric measurement.

6.6.1.3.1.2 Affordability

A major driver in automation is the perception that automated systems are less expensive to build and operate than manned systems. This is especially true in the case of small UMVs. Small UAVs such as the Raven and Pointer can be fielded inexpensively with non-pilot personnel trained to operate them, further reducing the

operational costs. Since the initial cost is low, these assets can be considered attritable. That is, loss of the asset does not result in catastrophic mission failure. This capability can, in turn, lead to a greater degree of mission flexibility, another advantage of UMVs, as mentioned above.

Operating costs are also often cited as an advantage of UAVs. This reduction in logistic footprint is realized since there is no need for pilots and supporting equipment. For example, 24 hour surveillance can be maintained by a single U.S. Army tactical UAV Shadow system, which includes 3 aircraft. This system has a logistics footprint of about 16 soldiers [9]. A similar capability from a manned aircraft system would require a much larger compliment.

6.6.1.3.1.3 Workload

Automation is often implemented to reduce operator workload. There have been some very important examples of this such as the crew reduction from three to two in commercial aircraft. The 1970s saw an increase in automation which reduced workload enough on commercial flight decks to lead to this crew reduction. However, it has also been shown that automation often changes the nature of workload. The automated cockpit, for example, changes workload from physical to cognitive. For an already busy operator, that shift can dramatically increase their overall workload. Attempts to lower workload through automation need to be carefully evaluated and empirically demonstrated.

There are many important benefits of automation as discussed above, however designers must take great care not to gain through automation in one area only to detract from another area. If not properly addressed, the "costs" of automation can be greater than the reward. Some potential costs of automation are discussed below.

6.6.1.3.2 Potential Costs of Automation

6.6.1.3.2.1 Mode Awareness

A major concern regarding human-automation interaction has to do with mode awareness. Too many examples are available that show crews taking an action that would be correct in one mode, but that leads to problems in the present mode [10]. Consider the tragic example of Korean Airlines 007. This flight from Anchorage, Alaska to Seoul, South Korea ended in a tragedy that was a confluence of many factors, but the initiating cause was mode awareness. The pilots were in heading hold mode which keeps the aircraft on the generally correct heading within 15 degrees. This is adequate for short distances, but as will be seen the error grows dangerously large over longer flights. The pilots switched the mode control panel from heading hold to the more accurate inertial navigation system (INS). However, entry into INS mode requires satisfaction of two conditions. The aircraft needs to be within 7.5 miles of the route and pointed in the general direction of the route. One of these conditions was not met. The aircraft, therefore, never entered INS mode. Over thousands of miles, the aircraft drifted 200 miles off course into airspace controlled by the USSR and the event ended with disastrous consequences. The initial cause of this accident was the pilots' misunderstanding of the control modes. Clearer enunciation and more intuitive control of the modes are essential to making the automation more useable and error tolerant.

6.6.1.3.2.2 Out of the Loop

In a highly automated system, the operator may be asked to monitor a number of processes. If the automation controlling these processes is largely successful and failure rates are low, the operator may not need to monitor very closely. This can lead to the operator being out of the loop when a failure does occur.

Recognizing that a failure has occurred and getting back the situation awareness needed for diagnosis can take a critically long time. This concern is especially prevalent for "semi-autonomous" systems, in which the system is highly automated without a human in the loop, until something goes wrong. The role of the human is then to quickly assess the situation and take corrective action, but having been out of the loop, this task is much more difficult. Lee and Moray called this "Out Of The Loop Un-Familiarity" [11].

6.6.1.3.2.3 Knowledge of Automation State

Highly related to being out of the loop is knowledge of automation state. In order for the operator to team with the automation, he/she needs to know what the automation is doing and why. The automation needs to be transparent and not a "black box." The lack of this knowledge will lead to mode awareness problems, under utilization, high workload in trying to determine what the automation is doing and poor situation awareness.

6.6.1.3.2.4 Over Reliance

Automation bias can come in two forms, over reliance on automation and under reliance. Over reliance, also called complacency, takes place when operators trust the automation to the extent that they no longer cross check what it is doing, and blindly accept its direction. An example of complacency was discussed by Azar [12]. A Panamanian cruise ship, Royal Majesty, was off the coast of Nantucket. The ship was being controlled by a satellite navigation system, which failed. Several other sources of navigation information were correct and available to the crew. These however, went unmonitored and as a result the ship ran aground.

6.6.1.3.2.5 Under Reliance

Automation can also be biased toward under reliance. High false alarm rates in the early design of the Ground Proximity Warning System (GPWS) led pilots to disable the system and turn off the automation. This has also been called automation disuse by Parasuraman and Riley [13].

6.6.1.3.2.6 Brittle Automation

Automation is brittle when it works only in specific situations and doesn't generalize well to unanticipated situations. This occurs when the models of the world instantiated in the automation are incomplete or inaccurate. This can cause users to lose confidence in the automation and not use it in situations where it may work well [14]. In these cases, it's important to provide a cooperative aid, rather than a fully automated decision or solution. In this way, the operator can provide the flexibility needed.

6.6.1.3.2.7 Accountability

An interesting aspect of higher levels of automation is that it is less and less clear who is accountable for an erroneous action. In a manned system the pilot or driver is in control and accountable for the actions of the vehicle. However, especially under higher levels of automation, it is not clear that the UMV operators are making the decisions that account for actions of the vehicle. While this operator is still legally responsible, he may be accepting decisions made by the automation. The decision logic may have been designed by the person who developed the automation algorithms or the software engineer that produced the code. This question becomes far less academic when UMVs become armed vehicles.

6.6.1.4 Philosophies and Methodologies

6.6.1.4.1 Human-Centered Automation

The introduction of glass cockpits and subsequently automation to piloted fixed wing commercial aircraft led several researchers to examine the transition to automated cockpits and the human automation interaction. Wiener and Curry [15] studied the transition from traditional steam gauge cockpits to glass cockpits. This study documented problems encountered during initial transition to automation. A major issue noted by Wiener [16] is that while automation may reduce small errors, it may invite large blunders, such as the American Airlines accident in 1996 [21]. Wiener (1989) [17] found that the present generation of automation was essentially sound, but lacking in proper user interface design.

Billings [18] synthesized much of the extant work on human automation. Billings defines human-centered automation as automation designed to work cooperatively with human operators in the pursuit of stated objectives. Building on previous work, Billings developed guidelines for human-centered automation, shown in Table 6-4.

Table 6-4: Human-Centered Automation Guidelines [18]

- The human operator must be in command.
- To command effectively, the human operator must be involved.
- To be involved, the human operator must be informed.
- The human operator must be able to monitor the automated systems.
- Automated systems must be predictable.
- The automated systems must also be able to monitor the human operator.
- Each element of the system must have knowledge of the others' intent.
- Functions should be automated only if there is a good reason for doing so.
- Automation should be designed to be simple to train, to learn, and to operate.

As can be seen from earlier examples, many of these guidelines have been violated in interface designs that led to predictable results. Guidelines can be useful in developing an underlying philosophy of design. Following them to specifically design an interface, however can be very difficult. Other approaches have emerged to address human-automation interaction.

6.6.1.4.2 Intelligent Entities/Delegation

Another way to view the interaction with UMVs characterizes them as intelligent entities and seeks to delegate authority to UMVs in a systematic way. The playbook approach [19] uses the sports analogy that the operator is the coach and can call a "play." The intelligent entities know what their roles and responsibilities are within that play and can execute those autonomously. For example, an operator may need to get information about a particular intersection of two roads. He/she would call the "recon a point" play. The UMV asset or assets know what this means, a) move into a clandestine position, b) observe the intersection for a specified amount of time and c) return the sensor information to the operator. These plays can be modified by the operator as he/she sees fit.

6.6.1.4.3 UMVs in a Network Centric System

Yet another way to view UMVs is as information sources. A commander may have a network of UMVs at his or her disposal. Instead of specifically commanding any asset, he/she would instead ask the network for information. For example, if the commander needs information about enemy assets in a certain area, he can ask the network for that information. The network knows what assets are available, and which are needed for this request. Based on the priority of the request, the network assigns assets to obtain the requested information and forward it back to the requesting commander. In this network centric application, the UMVs are efficiently utilized and available to a wider tactical community.

These approaches are not mutually exclusive and indeed may work well together. Further, additional ways to look at automation have also been suggested: automation as an associate [20], UMVs working as a team or swarming, and automation serving as an electronic crew-member or wingman. The task of the interface designer is to determine how these may be adapted to best serve the operator's needs. They may serve as a toolkit for human-automation interface design.

6.6.1.5 Interface Implications of Automation

This section summarizes some of the more salient points of the preceding discussions and draws on other sources to provide issues to consider in UMV interface design. They are not meant to be exhaustive guidelines or a checklist, but rather things to look out for and consider.

1) Automation does not reduce operator workload per se; it may change the nature of the workload or may even increase it. Automating the "stick and rudder" flying of a UAV or "hand control" of a UMV to way point control reduces the continuous in the loop manual workload. However, the operator is now a supervisor of this automated system and has to monitor the vehicle state and the automation controlling the vehicle. The cognitive workload associated with this supervisory control may well be higher than the workload of physical control.
2) It is critical for appropriate use of automation that the user understand how the automation works and what mode the automation is in. Without this understanding, automation bias can result.
3) There is an inexorable trade off between higher levels of automation and unpredictability. As systems move more and more toward automation, they are, by definition, less predictable and therefore don't allow predictability or insight into what the UMV is doing (see # 2).
4) All automation is not created equal. It can be brittle, unpredictable, and prone to bias. Knowing about these pitfalls is half the battle. A designer must carefully look at where and how the automation may fail. Is the mission at a critical point? Does the automation gracefully degrade? Does the operator know the automation is failing?
5) In UMV operation, there are two tasks: vehicle control and sensor control/interpretation, however, they are often treated as one. An earlier section discussed levels of automation, the first taxonomy is almost entirely in terms of vehicle control. The second taxonomy from the Army Science board addresses a bit more mission oriented tasks, but is none the less a unitary scale. It's quite likely that the sensor interpretation task is the more difficult of the two tasks, and the more difficult to automate. As programmatic goals move toward control of multiple UAVs and even swarms, it's important to ask who is interpreting the sensor information and/or if that task is automated.
6) Beware of systems described as "semi-autonomous." Many current systems are designed such that a person is the ultimate monitor and failsafe, but if automation levels are high, the operator is prone to being out of the loop (see above). It is dangerous to design systems that require humans to manage

contingencies when the automation fails. However, it should be noted that some excellent work has been done in this area under names of shared control, traded control and situation adaptive autonomy. The reader is encouraged to fully explore these areas when faced with "semi-autonomous" systems.

7) UMVs are tools to provide information. It should be remembered that UMVs are just one more way to provide information to commanders and decision makers. As such, it is important to focus on the information coming in and the information presentation.
8) Lessons learned from manned aircraft automation and other domains such as power plants offer invaluable lessons to draw from. Years of automation errors and poor implementations provide rich information sources.
9) Individual differences in the use of automation make predictions difficult (Parasurman and Riley [13]). It is important for the user to know the rationale and pay-offs for using the automation.

There is no formula for interfaces with automation. There are guidelines as have been discussed, but most importantly the designer needs to be aware of these issues, communicate to the user and continually evaluate the system.

6.6.1.6 Research Issues

In many respects, work with UMVs and human-automation interfaces is in its infancy and many research questions remain. A great deal of research is called for to help answer these questions. In the area of automation, here are a few key research areas. These research areas are also addressed in Chapter 7 of this volume.

6.6.1.6.1 Level of Automation

An early section discusses the various levels of automation and levels of control. However, it's not clear which of these or what combination of these are optimal. Should a designer always try for the highest level of automation that is possible? Increasing levels of automation will lead to more mission capabilities, but at what cost? How will the operator manage off-nominal events?

6.6.1.6.2 Level of Involvement

How can a designer increase levels of automation, but keep the operator informed and in control? Should the designer keep the automation level lower to keep the operator in the loop? Or could they build highly automated systems, but build in tasks that keep the operators' situation awareness high? What interfaces facilitate this?

6.6.1.6.3 Automation Transparency

How can automation interfaces be designed so that the actions and intentions of system are clear to the operator? Pilots often struggle with understanding what the flight management system is "trying to do." This needs to be avoided. The intent of the automation needs to be clear to the operator in control. How this can be done, to what extent it is necessary, and what interfaces facilitate this are major research questions.

6.6.1.6.4 Training

The question of operator qualifications have become very important, especially when in control of UAVs. Do the operators need to be qualified pilots or can they simply be trained on the system that they are flying.

These two different schools of thought have emerged, but there is little research available addressing this issue. A clear definition of the knowledge, skills and abilities to operate UMVs is needed to help answer this question.

There also remain many issues not raised in this chapter, but that will need to be addressed by UAV designers in the near future. The U.S. Army has plans to use manned helicopter and UAVs as teams. How should an interface be designed that optimises teaming between manned assets and unmanned assets? They also have plans to control (at some level) UAVs from helicopters. What is the optimal level of automation? Of control? What are the information requirements?

A last issue that will be raised here is one of commonality, or lack thereof. Manned aircraft have developed standards over the years, the classic example being the T arrangement of the airspeed, altitude and heading in an aircraft cockpit. This has allowed pilots to move from one aircraft to another with minimum levels of negative transfer. No such standards exist for UMV control station design. This has led to vastly different designs by each manufacturer and the result that operators must be trained very specifically on each platform control station, with little or no advantage of previous learning. This lack of standard design must be addressed for UMVs to reduce training costs, logistics and operation errors.

6.6.1.7 References for Implications of Automation/Autonomy

[1] Oxford Dictionary of Current English (1998). Oxford University Press: Oxford, New York.

[2] Sheridan, T.B. (2002). Humans and Automation: System Design and Research Issues. John Wiley & Sons, in cooperation with HFES: Santa Monica, CA.

[3] Sheridan, T.B. and Verplank, W.L. (1978). Human and Computer Control of Undersea Teleoperators (Man-Machine Systems Laboratory Report). Cambridge: MIT.

[4] Parasuraman, R., Sheridan, T.B. and Wickens, C.D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man and Cybernetics, 30(3), 286-297.

[5] Office of Secretary of Defense. (2005). Unmanned Aerial Vehicle Roadmap: 2005-2030. OSD: Washington, D.C.

[6] Army Science Board. (2004). Autonomy Levels for Unmanned Systems Workshop. Huntsville, AL.

[7] Standard Interfaces of UAV Control System (UCS) for NATO UAV Interoperability (2004). STANAG 4586(2).

[8] http://www.aerovironment.com/area-aircraft/unmanned.html

[9] Dowell, S.R. (2005). Shadow Logistics Footprint, personnel communications.

[10] Degani, A. (2003). Taming HAL: Designing interfaces beyond 2001. New York: Palgrave MacMillian.

[11] Lee, J. and Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. Ergonomics, 35, 1243-1270.

[12] Azar, B. (1998). Danger of automation: It makes us complacent. APA Monitor, 29(7).

[13] Parasuraman, R. and Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2).

[14] Layton, C., Smith, P. and McCoy, E. (1994). Design of a cooperative problem solving system for enroute flight planning: An empirical evaluation. Human Factors, 36(1), 94-119.

[15] Wiener, E.L. and Curry, R.E. (1980). Flight-deck automation: Promises and problems. NASA Technical Memorandum 81206. Moffett Field, CA.

[16] Wiener, E.L. (1985). Beyond the sterile cockpit. Human Factors, 27, 75-90.

[17] Wiener, E.L. (1989). Human Factors of Advanced Technology ("Glass Cockpit") Transport Aircraft. NASA Contract report NCC2-377.

[18] Billings, C.E. (1997). Aviation automation: The search for a human-centered approach. Malwah, NJ: Erlbaum.

[19] Miller, C.A., Goldman, R.P., Fubj, H.B., Wu, P. and Pate, B. (2004). A playbook approach to variable autonomy control: Application for control of multiple, heterogeneous unmanned air vehicles. American Helicopter Society 60th Annual Forum Proceedings, pp. 2146-2157, Alexandria, VA.

[20] Miller, C.A. and Riley, V. (1994). Achieving the associate relationship: Lessons learned from 10 years of research and design. 3rd International Conference on Human-Computer Teamwork: Cambridge, U.K.

[21] Aeronautica Civil of The Republic of Columbia (1996). Aircraft Accident Report: Controlled Flight Into Terrain, American Airlines Flight 965, Boeing 757-223, N651AA, Near Cali, Colombia, December 20 1995.

6.6.2 Control/Display Interfaces for Decision Support

Control and display interface design has always been a key concern for operator station development. However, future control stations that employ a single operator to control multiple UMVs will require controls and displays that not only enable conventional operator tasking, but also support supervisory tasks as well. These tasks include quick assessment, judgment and reaction to the appropriateness of the automation's changing plans and actions, as well as continually assessing the impact of these changes on overall mission objectives, priorities, etc. Moreover, as the number of controlled vehicles increase, it will be a challenge for the operator to maintain situation awareness through long periods of nominal operations interjected with short periods of time-sensitive contingency operations. The UMV interfaces need to effectively cue the operator's attention to critical information and support rapid task re-allocation. Thus, UMV interfaces must be tailored to increasing system autonomy and novel decision support systems.

Research is needed to support the design of situation assessment and decision support technologies that maximize flexible, fault-tolerant supervision of multiple intelligent semi-autonomous UMVs by a single operator. This technology area is more fully discussed elsewhere in this report.

Research must also specifically consider the design of tailored controls and displays for UMV decision support systems. Although often overlooked, control/display interfaces are critical to obtaining maximum benefit from a decision support system. If the controls and displays do not support the operator's decision making process, the operator's effective interaction with the intelligent automated systems will be compromised and the benefits of the automation will not be fully realized. In sum, effective decision support control/display interfaces are vital to the UMV operator's ability to make timely, accurate decisions.

This section summarizes a survey of decision support literature that sought to determine the degree to which controls and displays are considered in the design of previous decision support systems, identified key decision support interface types, and presented a notional classification system. The full survey is more fully described elsewhere [1].

6.6.2.1 Methodology for Decision Support Interface (DSI) Survey

Thirty-four DSI research papers were surveyed. The sampled literature included journal articles, conference proceedings, book chapters, and technical reports. To be included in the survey, two qualifications had to be met. First, the research had to involve a problem domain in which a human user was challenged by requirements to make quick and accurate decisions. Thus, the documentation did not need to address the UMV domain; there are a multitude of application environments that require decision support, the findings of which are potentially applicable to UMV supervisory control. Second, the research had to involve the evaluation of an interface concept – a decision support control and/or display was manipulated as an independent variable. Thus, this survey focused on empirical evaluations while steering away from papers that broadly addressed decision support issues, theory, and system design.

For each document in the survey, notes were made in multiple columns of a table [1]. Besides bibliographic information, there were seven columns to capture information on the nature of the decision support interface used in the research and the results of the evaluation. The following further describes the content noted in each column:

- DSI Category: the general grouping of the decision support interface (DSI) concept (e.g., attentional cue, status display, etc.).
- DSI Description: the purpose or intended benefits of the decision support concept.
- Intelligent?: the computational functionality of the DSI as to whether it is based on artificial intelligence algorithms, knowledge-based systems, modeling efforts, or simple rule-based logic.
- DSI Control Concepts: identifies any control concept used to interface to the decision support.
- DSI Display Concepts: describes the content/format of information presented in a DSI.
- Short Summary: an overview of the experimental design, independent variables and user tasks.
- Lessons Learned: summarizes results found with the particular decision support interface.
- Conclusions: lists DSI findings that may be generalizable to other decision support systems.

6.6.2.2 Decision Support Interfaces: Classification

Entries in the table were examined to derive major groupings. First, overall grouping of the decision support interfaces into ‘controls’ and ‘displays’ was accomplished. Next, further classification was accomplished based on other factors, such as the degree of computational functionality (intelligence) in implementing the interface and the nature of the support it afforded.

6.6.2.2.1 Decision Support Controls

Controls were first classified by control type (conventional, non-conventional). The majority of controls employed in the decision support interface literature sampled were conventional input devices – mouse and keyboard. Less often, researchers evaluated decision support control with novel devices. These included voice recognition, touch-screens, and reduced keyboards [2], and a haptic stick [3]. Within each control type, controls were grouped by technology type.

It was also deemed useful to classify by the method used to achieve control. For instance, researchers make the distinction between controls that are 'bottom-up' as opposed to 'top-down' [4]. In bottom-up control, individual elements of the decision problem are specified in detail, and the user then receives updates on each element. In top-down control, an overview of the decision problem is formulated, each element of the problem is then refined, and the user is guided down through the hierarchy of information. Another sub-classification is whether the control employed intelligent filters that retrieve, fuse, and manage information during the control process [3]. In contrast to having intelligent filters, the design can be based on the user changing system constraints, parameters, and/or plans which results in the initialization of processes that change (i.e., control) system states [2]. Two opposing control methods make up a final defined control category: closed loop control versus open loop control. With open-loop control, no feedback is sent back to the controller because the control equation is well-accepted. During closed-loop control, feedback is provided to the controller so that any subsequent input brings the system closer to the goal state.

6.6.2.2.2 Decision Support Displays

The sampled decision support interface literature described many types of displays. These were grouped into two major categories: displays that are simply rule-based and ones that are knowledge-based.

Automated warnings, alerts, and cues consisting of simple notifications (visual, aural, tactile) when a rule or goal state is broken (or met) constituted one common type of rule-based decision support interface display. These displays can be multi-modal. Attention cuing displays aid target prosecution and highlight critical information or points of interest through displays technologies such as synthetic vision symbology overlaid conformal on a camera video display, automatic target cueing, and assisted target recognition [5,6,7]. Status displays also mainly use a rule-base to present the decision maker with key parameters and their associated values. These displays can provide system history in the form of graphs, intelligence reports, or logs. Similarly, status displays can offer classification, filtering, or generalizing of situations and systems [8,9]. Action recommendation displays take status display information and use another rule-base to recommend a course of action to the decision maker (e.g., options, strategies), but does not implement the action [10,11]. A specific example is a Highway-in-the-Sky display that provides flight guidance [12]. Adaptive/adaptable systems are another type of rule-based display. Adaptable systems allow the operator to initiate automated state and mode changes, while adaptive systems allow the operator or automation to initiate the changes depending on the automation management schema. In some designs, a simple rule-base can cause automated mode changes when the operator crosses threshold values in path deviations, EEG index (workload measure based on electroencephalographic signals), or altitude (such as a ground collision avoidance system). Rules could also trigger the system automation to collect, filter, organize, and present information to the operator in anticipation of upcoming decisions.

Knowledge-based decision support displays are also referred to as intelligent displays. Definitions of "intelligence" vary, but many include the ability of an entity to achieve goals in complex real world situations. Some definitions specify that the entity must perceive, learn, reason, communicate, and act [13]. Using intelligent systems to support decision making is most appropriate for well-structured, clearly defined decision situations. Many types of intelligent displays rely on intelligent agents that work in the background to collect, filter, organize, and present information to the decision maker. With an intelligent system, status displays can also provide a feed-forward presentation to support the decision maker in anticipating future system states through Gantt charts and other predictive displays. Intelligent agents can also expand the capabilities of displays in adaptive systems. Other types of knowledge-based decision support displays are plan generators and evaluators. With these displays, operators can a) generate a plan and then have the system evaluate (critique) it; b) tweak a system-generated plan and then have it evaluated; or c) use the intelligent

technology to predict future system states by inputting "what if?" scenarios. The generation and evaluation of these plans are usually dependent on optimization algorithms or knowledge-based expert system modeling. Similarly, intelligent adaptive decision support displays rely on intelligent agents to monitor situations and decide what information to present to the user, in what format, and at what time. Some intelligent agents may replace some of the operators in a team decision making environment. To make effective team decisions, the agent-user collaboration, negotiation, and dialogue must work successfully so the agent can provide aggregate, inferential, and decision information to the operator(s) without information overload.

6.6.2.3 Decision Support Interface Survey: Lessons Learned

6.6.2.3.1 General Findings

One objective of the survey was to determine the degree to which controls and displays are considered in the design of decision support systems. For the literature sampled, the majority dealt with display interfaces as opposed to control interfaces. This finding is not surprising and can be viewed as reflecting the larger body of research dealing with displays, compared to controls, in overall human factors literature.

Another finding raises a more pressing concern. First, note that all the sampled literature focused on evaluating a control and/or display interface for a decision support system. However, less than a third explained the rationale for the control/display concept chosen to be employed with the particular decision support system. In other words, these reports didn't cite published research or a previous application of the control/display that supported its utility in the decision support system featured in the report. Although this finding may only reflect incomplete documentation by the authors, it is feared that it is indicative of an inadequate prioritization on control/display design. Another possibility is that there are problems inherent in generalizing existing control/display findings to decision support systems. Regardless, this finding supports the need for additional research on how best to apply control/display interfaces for use with decision support systems for UMV operators.

Below is a summary of survey findings. Complete results can be found in [1].

6.6.2.3.2 Decision Support Controls

The literature survey indicated that the use of non-conventional controls, such as speech recognition, touch-screens, and reduced (custom function) keyboards generally improved operator performance and either reduced or had no negative effect on perceived workload [2]. Another finding was that in many decision support systems, there are alternate control methods simultaneously available to the operator. For instance, the decision support system may allow the operator to choose:

- Information search/acquisition method, such as bottom-up (start with detailed query and get updates/ notifications) or top-down (global perspective with guides to detailed information);
- Information assessment method (what filters or constraints are employed);
- Whether to view recommended alternatives or only the "optimal" solution;
- Control input device (by sketching, typing, or speaking novel problem solution into the system).

Empirical results suggest that these control choices could significantly affect the decision support effectiveness and overall mission performance. For instance, choosing to select a path recommended by the decision support system, as opposed to using a more cumbersome sketch/graphical input method, can bias the operator's decision [9].

6.6.2.3.3 Decision Support Displays

6.6.2.3.3.1 Rule-Based Displays

Automated Warnings, Alerts, and Cues

Similar to decision support controls, performance with non-conventional decision support displays was generally better than with conventional visual displays. Multi-modal displays were found to improve performance for many tasks, especially when used as a redundant cue to visual information. Examples include haptic [3], auditory [14], and voice displays/feedback [2,15]. Another key finding related to cuing displays is the potential for cognitive tunneling. Cognitive tunneling can occur when the operator becomes focused on the cue being displayed to such an extent that other important objects in the view are not attended [16]. Research is needed to explore how cue displays might be designed to minimize the negative effects of cognitive tunneling. Decision support cues can also unintentionally obscure both expected and unexpected information [7]. Techniques are being explored to minimize this problem [17]. Finally, there is a possibility that operators will over rely on cues displayed from the decision support system, resulting in operator complacency [15].

Status and Action Recommendation Displays

In general, the sampled research showed that status and action recommendation type decision support displays were beneficial to operators' performance in the task environments tested. However, the reliability of information displayed begins to have a more consequential impact with recommendation displays. This is because it is more difficult for operators to judge from a recommendation display that underlying data is suspect. Studies have shown that when the decision support system has reduced reliability, operator performance is better with a status display by itself [10,18]. Additional findings suggest displays that incorporate graphics are better than text-based displays. It was hypothesized that the graphic displays were rated higher in usability because they support human pattern recognition or recognition-primed decision-making [8].

6.6.2.3.3.2 Knowledge-Based (Intelligent) Support

Since knowledge-based displays are more difficult and time-consuming to develop and require a well-defined domain, empirical testing is scarce and this was reflected in finding only a few relevant documents in the literature sampled. One evaluation of an advisory tool designed to generate flight paths around weather showed that providing a display to aid operators to manually develop their own plan prior to getting a plan from intelligent agents enabled operators to catch faulty plans better compared to displays that only provided an auto-generated plan [19]. This result suggests that a cooperative process between the operator and the decision support system would help minimize the occurrence of automation bias.

The nature of the cooperative process can also impact the effectiveness of the decision support system. One study [20] demonstrated that operators could generate plans for distribution to team members faster with an intelligent planner that provided displays to assist the operator through the planning process (providing suggestions, blanks to fill, etc.), than with an intelligent planner that critiqued an operator-generated plan (displaying how good it was according to critical mission measures). This illustrates how display design needs to simultaneously take into account control design as well. In other words, the optimal nature of the operator-system dialogue for capitalizing on the benefits of the decision support system needs to first be determined and then supported by the control/display interface.

Information reliability also plays a role in the effectiveness of decision support system displays. In the study mentioned above [20], a display with no support was compared to three different types of decision support

displays: one that just listed raw data, one with the data classified into columns, and one in which accuracy probability information was also provided. By manipulating the source of errors (raw data, erroneous classification, erroneous probability), the results suggest that presenting the data classified into columns was best, despite the fact that subjective ratings indicated that operators trusted all three types of displays equally. Also, the finding that operators' trust was the same for all display types illustrates the value of presenting key information in the display by which the operator can judge the adequacy of the data and, ultimately, improve the decision making process.

As an example of a promising knowledge-based decision support display, consider [21]. These researchers developed a tactical display concept consisting of a set of PCE (Predicted-Capability-Envelope) contours overlaid on a 2D map representation (Figure 6-31). These PCE contours describe the maneuvering margins of ownship movements, in relation to other moving platforms' motions and ownship capabilities. Capability prediction is a model-based concept that takes dependencies between control actions and effectiveness parameters into account. This involves calculation and presentation on a navigation display of the total maneuvering margins, the so-called PCE. The concept was developed for a ship-maneuvering environment. In this case, the predicted margins include restrictions due to boundaries and other traffic ships. The PCE represents the complete reach of the controlled vessel for a particular time horizon. By intersecting the PCE with a required minimum safety distance, an integral representation of (other traffic) threats and (controlled ship) capabilities is obtained. Provided that course and speed of other ships remain the same, the presented threats will be geographically stable areas. Navigators may consider these threats as obstacles in the fairway. Thus, an integrated navigation display is obtained which provides an overview of the ship's maneuvering and collision avoidance information for a particular navigation task. With respect to the PCE concept, several in-depth part-task simulator studies were performed to quantify the effects of using PCE on navigator performance. The experimental results show that capability prediction is far more effective than other prediction information, e.g., path prediction [22].

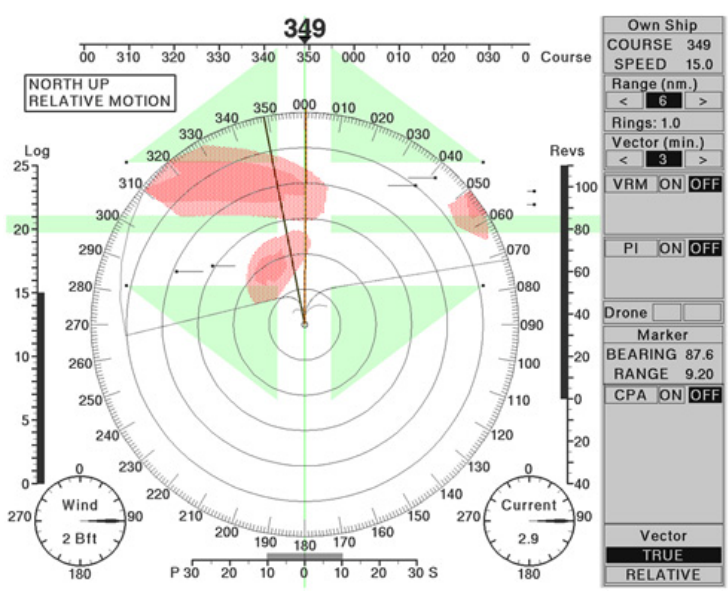


Figure 6-31: PCE Display Representation as an Overlay on a Radar Display. The center dot represents ownship position. Starting at this position, an area is marked (thin grey envelope) for which proximity information is calculated. The red zones represent areas in which the own ship will be within one mile passing distance from another vessel.

6.6.2.4 Survey Implications on UMV Supervisory Control

Decision support interfaces are key to realizing the benefits of intelligent automated systems that facilitate single operator control of multiple UMVs. If the interface fails to support the operator's decision making process, the operator's ability to make timely, accurate decisions will be compromised as will the ability to judge the operation of the semi-autonomous systems. This literature survey, however, suggests that attention to date has focused on development of the knowledge-base for the decision support rather than the optimal design of the controls and displays used with the decision support system. Thus, there is a need for research that systematically addresses interface issues for various categories of UMV decision support systems.

Some directions for this research are indicated in the surveyed literature. For instance, research on multi-modal interfaces (e.g., speech-based input and visual/aural redundant displays) supports their use with decision support systems. The interaction of control and display design was also illustrated, showing that the desired operator/system dialogue or collaboration must also drive interface design. Other studies demonstrated the value of presenting key information in the display by which the operator can judge the adequacy of the data and, ultimately, improve the decision making process. For UMVs, this would suggest displaying status information for each UMV under supervision, increasing the displayed information by a factor of the number of vehicles. To minimize information overload/retrieval problems, an intelligent system which highlights important information while maintaining the availability of other information may be useful, as well as innovative display approaches that support anticipatory decision-making.

In research evaluating candidate UMV controls and displays, the reliability, bandwidth, and timeliness of information need to be manipulated to determine their effect on the utility of the decision support interface. It is also important that the interface design/evaluation takes into account issues of cognitive tunneling, effective information retrieval, and automation bias. Workload and vigilance demands are also critical.

6.6.2.5 References for Decision Support Control/Display Interfaces

[1] Calhoun, G., Ruff, H., Nelson, J. and Draper, M. (2005). Survey of decision support control/display concepts: classification, lessons learned, and application to unmanned aerial vehicle supervisory control, Proceedings of the 11th International Conference on Human Computer Interaction, Las Vegas, CD-ROM.

[2] Winter, H., Champigneux, G., Reising, J. and Strohal, M. (1997). Intelligent decision aids for human operators. AGARD Symposium on Future Aerospace Technology in the Service of the Alliance, AGARD-CP-600, Vol. 2, Palaiseau, France, 14-17 April.

[3] Scerbo, M.W. (2000). Adaptable and adaptive technology in the lab and in the field. Proceedings of the Human Performance and Situation Awareness in Automation Conference, pp. 57-62.

[4] Sycara, K. and Lewis, M. (2002). From data to actionable knowledge and decision. Proceedings of the Fifth International Conference on Information Fusion, Annapolis, MD, July 7-11.

[5] Endsley, M.R. and Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 42(3), 462-492.

[6] Hughes, S. and Lewis, M. (2002a). Attentive interaction techniques for searching virtual environments. Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD., pp. 2159-2163.

[7] Hughes, S. and Lewis, M. (2002b). Directing attention to open scenes. Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD., pp. 1609-1612.

[8] Morrison, J.G., Kelly, J.T., Moore, R.A. and Hutchins, S.G. (1998). Implications of decision making research for decision support and displays. In: J.A. Cannon-Bowers and E. Salas (Eds.), Making decisions under stress: Implications for individual and team training. Washington, DC: APA Press, pp. 375-408.

[9] Smith, P.J., McCoy, C.E. and Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: The effects on user performance. IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, (27) 3, May, pp. 360-371.

[10] Crocoll, W.M. and Coury, B.G. (1990). Status or recommendation: Selecting the type of information for decision aiding. Proceedings of the Human Factors Society 34th Annual Meeting, pp. 1524-1528.

[11] Woods, D.D. (1986). Paradigms for intelligent decision support. In: Holnagel, M., and Woods, D. (Eds.). Intelligent Decision Support in Process Environments. New York: Springer-Verlag, 153-173.

[12] Snow, M.P. and Reising, J.M. (1999). Effect of pathway-in-the-sky and synthetic terrain imagery on situation awareness in a simulated low-level ingress scenario. Proceedings 4th Annual Symposium on Situational Awareness in the Tactical Air Environment, Piney Point, pp. 198-207.

[13] McCarthy, J. (2003). What is artificial intelligence? Stanford University, web essay, from http://www-formal.stanford.edu/jmc/whatisai/whatisai.html

[14] Olson, W.A. and Sarter, N.B. (2001). Management by consent in human-machine systems: When and why it breaks down. Human Factors, (43)2, pp. 255-266.

[15] Foy, L. and McGuinness, B. (2000). Implications of cockpit automation for crew situational awareness. Proceedings of the Human Performance and Situation Awareness in Automation Conference, pp. 101-106.

[16] Yeh, M. and Wickens, C.D. (2001). Display signaling in augmented reality: Effects of cue reliability and image realism on attention allocation and trust calibration. Human Factors, 43(3), 355-365.

[17] St. John, M., Maines, D.I., Smallman, H.S., Feher, B.A. and Morrison, J.G. (2004). Heuristic automation for decluttering tactical displays. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, HFES: Santa Monica, CA, pp. 416-420.

[18] Cummings, M.L. and Guerlain, S. (2005, in review). The decision ladder as an automation planning tool. Cognition, Technology, and Work, Springer-Verlag London Ltd.

[19] Guerlain, S. (2000). Interactive advisory systems. Proceedings of the Human Performance and Situation Awareness in Automation Conference, pp. 166-171.

[20] Lewis, M. (1999). Anticipation delegation, and demonstration: Why talking to agents is hard, In: M. Klusch, O.M. Shehory, and G. Weiss (Eds.), Cooperative Information Agents III, Lecture Notes in Computer Science, Vol. 1652, Springer-Verlag Berlin, Heidelberg, pp. 365-389.

[21] van Breda, L. and Passenier, P.O. (1998). The effect of path prediction on navigational performance. The Journal of Navigation, 2, 216 228. Cambridge, UK: Royal Institute of Navigation.

[22] van Breda, L. (1999). Anticipating behaviour in supervisory vehicle control. PhD thesis. Delft, The Netherlands: Delft University Press.

6.7 ADDITIONAL CONSIDERATIONS FOR UMV INTERFACES

The environment and methods in which a UMV is intended to operate will dictate particular UMV's control station affordances and constraints. This section examines the scale and variability of control stations for UMVs, identifying primary affordances and unique constraints associated with each. This section also discusses issues associated with controlling UMVs from moving platforms.

6.7.1 Scale of UMV Operator Interface

Scale, in this context, is used to describe the constraints imposed upon the design of UMV control stations. The scale ranges from man portable to large space platforms and all are limited to single operator control.

6.7.1.1 Man-Portable Platforms

In man portable platforms, the UMV is ideally very small, lightweight, rugged, and easy to operate [1]. It is essential that this platform meet these characteristics because a typical operator will be transporting the entire UMV system along with other mission critical equipment. These UMVs are usually used for "what's over the hill" type missions; requiring fairly autonomous operation to gain time critical, nearby information. The control stations for man portable UMVs will typically be no larger than a laptop and can be made to fit on smaller devices such as a PDA or head-mounted displays. For example, the Pointer UAV, developed by AeroVironment Corporation, is operated by the user through a large tablet-like PDA [2] (Figure 6-32). Another example of this can be seen in [3]. However, ruggedization of this equipment will generally increase weight and size.

Figure 6-32: Example of Man-Portable Control Station. (Source: http://www.aerovironment.com/).

6.7.1.1.1 Affordances Provided by Man-Portable Platforms

Of the three main classes of UMV control stations, man portable stations afford the fewest resources for interfacing with the vehicle. Often times man portable UMVs require direct line-of-sight control of the vehicle. This allows more immediate responses to control inputs. The short duration of operations and lack of control station equipment should minimize the amount of ergonomic concerns with respect to body posture, workstation design, etc. Decreased available interface real estate requires simplified control inputs into the UMV system, and varied operating environments may allow for more unconventional input methods into the system such as speech recognition.

6.7.1.1.2 Constraints Associated with Man-Portable Platforms

Interface designs must be optimized for essential basic functions, leaving little room for displays and controls that may expand capability. Further, the operating environment of a man portable UMV control station can vary more than restricted space and unlimited space platforms, increasing the need to make robust, hardy equipment. Control input devices cannot be overly sensitive or delicate (e.g., a PDA-style stylus may be too fragile of an input device if the operating environment requires all weather operations). Display screens must be able to be viewed under less-than-ideal conditions (e.g., bright desert environments), and access to Command and Control (C2) information for these systems is limited. Operators must act as the vehicle director and perform multiple other functions that need to be executed, requiring a large amount of autonomy on the vehicle's part. Man portable UMV platforms create many challenges if an operator needs to operate more than one vehicle at a time.

6.7.1.2 Restricted-Space Platforms

Restricted space control stations are characteristically built to be semi-mobile, often in a portable trailer. However, restricted-space platforms may also be in the back of another vehicle, such as a High-Mobility Multi-purpose Wheeled Vehicle (HMMWV) or tight quarters on a ship. Space is often limited, varying based on the UMV system it supports, but there is no requirement for the system to be man portable. An example of a restricted space control station can be seen in Figure 6-33.

Figure 6-33: Example of Restricted-Space Platform.

6.7.1.2.1 Affordances Provided by Restricted-Space Platforms

UMV restricted space control stations possess much more interface space compared to man portable systems. Additionally, the UMV control station operating environment can be managed to a large degree with respect to the station's temperature, lighting, noise, etc. This allows some degree of flexibility in the design of the displays and controls, as they do not have to be as hardy. The increased space in these stations also allows operators to have additional, but not necessarily vital tools, to improve areas such as situation awareness and workload. However, it is important to remember that display and information input space, while larger the man portable systems, can quickly become congested. Another benefit to the restricted space control station is the greatly improved access to C2 information as these platforms have the potential for SATCOM communications.

6.7.1.2.2 Constraints Associated with Restricted-Space Platforms

As would be expected, restricted space platforms have fewer constraints than man portable stations, but more than unlimited space designs. Often, achieving a good ergonomic layout in restricted space control stations is difficult. Vehicles operating from these control stations usually demand high levels of operator oversight, requiring multiple display and control surfaces. Fitting all the controls and displays is typically the first priority; accommodating the human is secondary. Maintenance access is often restricted as is any redundancy in the controls or displays. Information input is often performed through the use of conventional mouse, keyboard, and/or joystick devices. Technologies such as speech recognition can ease the space requirements for control surfaces in these platforms and allow for more display area, better ergonomics, and increased capability tools. Human interface design needs to consider how best to access and display large amounts of information on few displays (i.e., what is an appropriate cognitive distance for each item, how much digging for information should be required – information depth versus breadth, etc.).

6.7.1.3 Large-Space Platforms

In a platform where space is abundant and more than able to accommodate the operators, the station would most likely be located in some sort of bunker or building [1]. These arrangements allow for the greatest degree of control with respect to the control station operating environment (e.g., temperature, lighting, noise, etc.), but the least amount of mobility. These systems might typically be large platforms with highly autonomous vehicles performing numerous functions that are all tied into net centric feeds. Some potential examples of large-space platforms can be seen in research with data walls, as shown in Figure 6-34.

Figure 6-34: Example of Large-Space Platform – A Data Wall at U.S. Air Force Research Laboratory.

6.7.1.3.1 Affordances Provided by Large-Space Platforms

Perhaps the most appealing aspect of a large space platform is that it could allow for a truly user-centered design approach as defined by [4]. While the user-centered design approach should be used for the man portable, restricted space, and large space platforms, it can more fully be exploited in the large space system. Using this approach, the system could be designed with the user in mind at all stages, maximizing the technologies that would allow an operator to function at peak efficiency and effectiveness. It is possible to incorporate all expanded capabilities as well as basic functionality. The size could be enlarged to fully accommodate proper ergonomic design, information input, and display space needs of a human operator. In the large space platform, it is possible to completely control the operator's working environment – ensuring no stray lights, sounds, or other disturbances interfere. This type of platform would be ideal for controlling multiple UMVs or any type of system that requires crews of operators.

6.7.1.3.2 Constraints Associated with Large-Space Platforms

The unlimited space design is not without constraints. The biggest drawback is its lack of mobility. To move the entire station from one area to another becomes a long and difficult task. There is also an issue of oversaturation. For instance, just because there may be space to add an additional monitor to the workstation does not mean it would improve the operator's performance. Providing an operator with additional channels of information to monitor can decrease overall situation awareness, increase workload, increase the possibility of errors due to missed information or misperceived information, increase the possibility of cognitive tunnelling, as well as a host of other problems. Human factors engineering would need to ensure the system does not inundate the user with too much information or too many controls.

6.7.1.4 References for Scale of UMV Operator Interface

[1] U.S. Department of Defense. Unmanned Aircraft Systems Roadmap 2005. Retrieved November 30, 2005, from http://www.acq.osd.mil/usd/Roadmap%20Final2.pdf

[2] AeroVironment, Inc. Pointer FQM-151A: Unmanned Aerial Vehicle (UAV) System. Retrieved November 3, 2005, from http://www.aerovironment.com/area-aircraft/prod-serv/ptrdes.pdf

[3] Goodrich, M. and Quigley, M. (2004). Mini-UAV telemetry and imaging visualization for searching tasks [Abstract]. Proceedings of the Cognitive Engineering Research Institute Annual Human Factors of UAVs Workshop, 1. Retrieved December 1, 2005, from http://www.cerici.org/workshop/abstract/Goodrich.pdf

[4] Norman, D.A. and Draper, S.W. (Eds.). (1986). *User centered system design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

6.7.2 UMV Operation from Moving Platforms

An interesting and even more complex problem arises when considering the control of UMVs from moving platforms. One possibility of this is a man portable system being controlled as the operator is on the run. The operator, now the moving control platform, must operate a UMV while actively maneuvering his/her own body. An alternative scenario involves a UMV operator controlling the UMV from a station located in a moving vehicle. Operators must maintain spatial orientation and awareness of the UMV, while simultaneously sensing and understanding the dynamics of their own vehicle [1].

The lack of proprioceptive feedback inherent with UMV control can increase the workload for operators trying to maintain spatial orientation and awareness. Sensory information afforded an operator is drastically reduced and delivered almost entirely through the visual channel. Operators controlling UMVs from moving platforms receive sensation cues completely independent (and often contradictory) from the vehicle being remotely controlled. This problem is compounded by the fact that the operating environments of the vehicles may be different and could change during the duration of the operator's control of the UMV. For example, an operator controlling a UAV must maintain spatial orientation and awareness of the UAV, while possibly operating from an air, ground, sea, or underwater vehicle. The proprioceptive cues of the ground, sea, or underwater environments can vary greatly from those of the aerial environment and would compound the mismatch of cues, increasing the difficulty of controlling a UMV and potentially leading to motion sickness.

In the above scenarios, the operator is simply operating the UMV while another operator controls the platform containing the UMV control station. However, it is also plausible that a single operator will directly control a vehicle (or his/her own body motion) while also attempting to control one or more UMVs. This presents some unique challenges [1]. While it may reduce manpower and equipment needs, it may also require some control devices and display surfaces to act as inputs and displays for both the manned and unmanned vehicles. Issues emerge as to how best to switch the control and display between the vehicles, and which input devices and display surfaces can be used for both vehicles and which should be solely dedicated. This becomes an issue of supervisory control – to what degree does the operator need direct control and what areas of operation should be made autonomous?

Regardless of who is controlling the operator's vehicle, certain key concerns have been identified by [2] that must be addressed for UMV operations from a moving platform:

- How should spatial information about own-vehicle and controlled vehicle(s) be displayed to the UMV operator?
- Can individual differences in spatial orientation and mental rotation be used as criteria for selecting and assigning UMV operators?
- What types of training and simulation systems are necessary to develop the necessary skills to manage complex multi-vehicle dynamics?

In addition to the challenges of navigation/spatial orientation, other issues of concern are motion sickness, biodynamic interference with manual control, and head-mounted display (HMD) bounce. Motion sickness can occur when the visual cues of motion differ from those perceived by the inner ear, creating a sensory conflict [3]. Motion sickness can include nausea, dizziness, disorientation, and increased stomach awareness. Another issue is that of biodynamic interference, or, the input to manual control devices from the shock and vibration of rough terrain, transmitted from the vehicle to the operator and into the controls. This is especially common in UGV operations. And finally, HMD bounce refers to the phenomena of HMDs resonating at certain frequencies of vertical axis vibration, which are often found in ground vehicles [4]. The HMDs "bounce" relative to the face and eyes of an operator, degrading visual performance.

Below is a general discussion of the physical ergonomic aspects of mobile (and fixed) UAV control stations. This is followed by a discussion of a few of the specific issues surrounding the control of UAVs from an air-based control station.

6.7.2.1 Physical Ergonomic Aspects

This section discusses physical ergonomic aspects to be considered to optimize working conditions for mobile and fixed UAV workplaces based on lessons learned in the design and testing of workplaces. The discussion

will give insight in important issues in a general sense; however, detailed solutions will not be given. These details can be found in applicable handbooks, literature, reports, and such. Most of the issues mentioned below may be recognized as common sense. However, recent experience has shown that even common sense is commonly overlooked.

The physical ergonomic aspects include: anthropometry, biomechanics, climate control, workplace lighting, vibrations/accelerations, and control and display placement. The *fixed workplaces* are defined as shelters/ containers with one or more UAV (control) workplaces. These shelters/containers are usually fixed on a certain location. However, it is also possible that these workstations will be operational during displacements; hence, they become *mobile workplaces*. The *portable workplaces* described here are in effect soldiers equipped with devices to assess information coming from the UAVs or even to control the UAVs.

6.7.2.1.1 Fixed and Mobile Workstations

6.7.2.1.1.1 Anthropometry

A common experience, at TNO Defence, Security & Safety, is that ergonomic specialists are asked to test the anthropometrics of a workplace once the design is more or less frozen without any prior involvement during the already fulfilled design process. The outcome of these tests is usually disappointing because the physical human operator is not considered properly: the operators do not fit, tall operators have to be shoehorned in and small operators can not reach applicable controls and displays.

Even more, the specialist's remarks may be such that an easy improvement of the workplace tested is not possible, invoking the need for drastic and costly design changes to the workplace. The most common cause for all this is that anthropometrics have not been considered properly during design activities despite commonly available human modelling techniques (see Figure 6-35). The most basic approach to overcome this issue is to set clear and proper anthropometric requirements once starting a new project, to assess these anthropometric requirements once developed and finally, to test the actual result using digital modelling, prototypes and/or mock-ups.

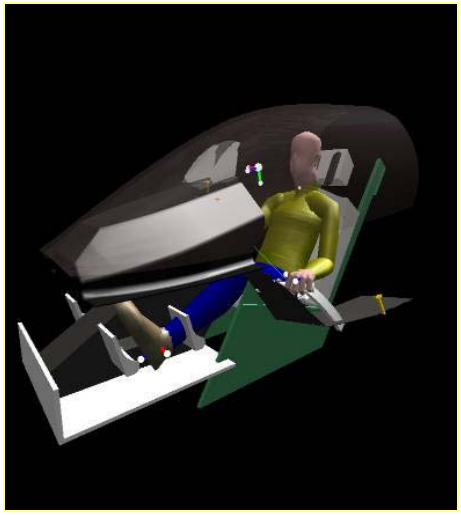


Figure 6-35: An Example of a Digital Human Modelling System: A Manikin in a CAD Model of the F16.

Today, the use of statistical boundaries, or percentiles, is being replaced by "cases". These cases define the boundaries based on manikins that have to be accommodated. The advantage of cases that they are more accurate, give a better accommodation result and they are easily embedded into currently used human modeling techniques.

6.7.2.1.1.2 Biomechanics

Frequently, the following question is put forward: what is the amount of force that a P5 female can exert? The question implies, in most cases, that small females are weak. However, it must be noted that size/stature and force are not correlated at all! There are small females that can exert more force than their tall male colleagues and vice versa. On the other hand, it must be noted that the reach capabilities of smaller persons are more limited, possibly affecting the amount of force to be exerted.

Biomechanics come into play for UAV control stations when thinking about antennas that have to be erected and while deploying a launch bed for the UAVs themselves. These riggings usually require a lot of energy before they are effectively deployed. The amount of energy and forces needed must be in accordance with the user population capabilities. More often, the situation occurs that there is a mismatch. Again, there is a basic approach to prevent mismatches taking place: set biomechanic requirements, assess these issues when developing, and finally qualify the design in field tests on prototypes/mock-ups.

6.7.2.1.1.3 Climate Control

Climatic requirements are usually well defined and tested. For instance, an entire shelter including its equipment is tested in a climatic room under certain conditions (e.g., A1, hot dry conditions, STANAG 2895 [8]) with a positive result. However, in the actual deployment, under A1 conditions it becomes clear that the required temperature cannot be reached: in effect it is too hot. The mistake, often experienced at TNO Defence, Security & Safety, is that the test settings do not represent the actual working conditions because the energy dissipated by the operators (about 1.3 kW per person) is not taken into account. The solution is to be very critical for climatic test, and to make sure that the testing conditions do represent actual working conditions, using all equipment, using all heat producing elements, etc.

6.7.2.1.1.4 Lighting and Colours

Displays are commonly used in today's workplaces. Blackout lighting (blue or red light) systems are also commonly used in control stations to cope with certain tactical conditions. One must make sure that the information shown on the displays is in accordance with the blackout lighting conditions. Recent experience, at TNO Defence, Security & Safety, showed that friendly troops, displayed in green, were not recognizable under red light conditions. Other elements, like enemy front lines, even disappeared under these conditions. Another problem became apparent after a short while: certain operators (about 8 to 10% of all males) could not see certain elements at all. These operators were unable to distinguish certain colors: they were colour blind. Therefore, a strong directive: take tactical lighting conditions and colour blindness into account in the design of displayed information in order to prevent any problems.

6.7.2.1.1.5 Mobile Workplaces

Designing workplaces for static conditions poses enough challenges already before an acceptable compromise is found – but an even bigger challenge becomes apparent when the workplaces have to be used during displacements as well. Aspects such as HIC (head injury criteria) and legislation for safety belts come into play. This section will not discuss all the rules, but will discuss an elementary issue for the operator. It is clear

that safety belts are elementary for mobile workplaces. The actual choice for a retractable or fixed safety belt has a strong influence on possibility to work effectively. Recent experience, at TNO Defence, Security & Safety, confronted operators with fixed belts disabling them to reach the controls and to read any display information: they were fixed to their seats and could not reach any workplace element. Therefore, a strong word of advice: opt at first for a retractable safety belt and make sure that the operator is not hampered while working due to any safety system.

6.7.2.1.1.6 Nuclear, Biological and Chemical (NBC) Conditions

It is a fact that our troops are confronted with NBC conditions. Even under these circumstances one must be able, when wearing protective gear, to effectively operate UAVs. It is clear that protective gear, especially the NBC gloves, affect dexterity. And the amount of energy that can be exerted is affected as well due to the isolating properties of the suit and the limited breathing possibilities when wearing a mask. All these elements combined will have to be taken into account for all operations inside *and outside* (e.g., when rigging antenna arrays, setting the UAV launch station) the UAV control station.

6.7.2.1.2 Portable UAV Workstations

There seems to be an apparent need for field operators (e.g., soldiers, policemen, firefighters), to have a mobile office. In some cases, the operator is simply equipped with an advanced mobile telephone. In other cases they have to carry a complete laptop while working. The result is, in most cases, not fully satisfactory: the equipment does not perform properly (e.g., one cannot access the worldwide web properly using a mobile phone due to the limited size of the display and input media) or it is simply too bulky (e.g., a complete laptop, including power source, is not compatible with standard battlefield operations). Recently, a system was developed as a prototype system, by TNO Defence, Security & Safety, with a helmet integrated vision device and a small touchpad as input device, in the Soldier Modernization Program (see Figure 6-36). This system basically focuses on the needs for the operator in the battlefield: some operators only need a display for a short period of time (e.g., a soldier needing information concerning his/her whereabouts), other operators need more information (e.g., a group commander needs more information and has an additional need to input data).

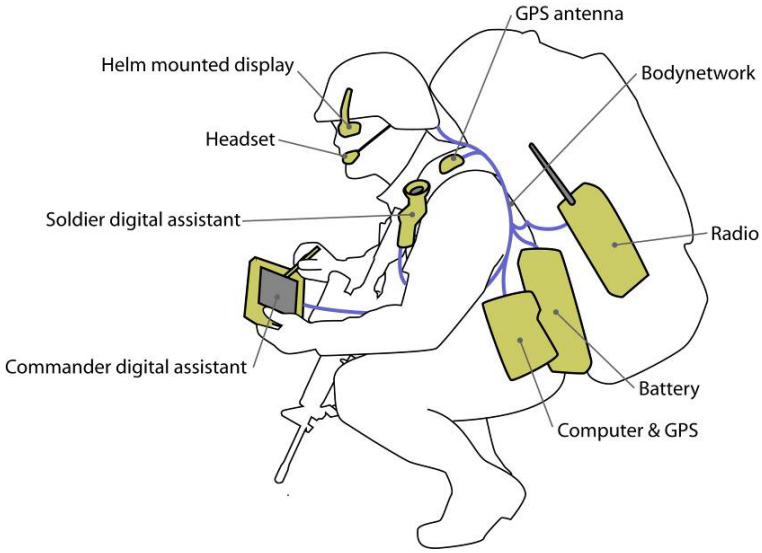


Figure 6-36: A Soldier Equipped with Various Portable Systems for Advanced Field Operations.

The basic thought behind this system is that it is fine-tuned to the operator's needs. This system is not yet used or tested in operational circumstances; this will take place in the beginning of 2006. An elementary issue to be focused on is the effect of the new system on battlefield operations. It is already apparent that some users will have to change their standard operational procedures because they will have an additional task to perform: the task of information management.

6.7.2.2 Manned-Unmanned UAV Teaming from Air-Based Control Station

Most currently used UAVs are controlled from a (closed) container Ground Control Station (GCS). The GCS operators are generally concerned with only one UAV system. Here, the Mission Commander communicates the acquired intelligence to other units. In past years, experimental setups have been developed in which the information from the UAV sensors is more directly presented to aircraft during a mission. A pioneer project was U.S. Army's AMUST (Airborne Manned Unmanned System Technology [5]). Here, an Apache helicopter and a Hunter UAV form a team, where the UAV may perform all kinds of useful sidekick tasks, such as reconnaissance, laser-designation of targets and acting as decoy. The UAV can be either controlled from a GCS or by the Apache's co-pilot/gunner.

The AMUST concept adds to the complexity of maintaining situation awareness: the operator (the co-pilot/ gunner) has to deal with at least three spatial frames of reference, namely the world, the helicopter and the UAV. Our first research questions were: Is such a co-pilot able to build up situation awareness involving multiple platforms, and how is this situation awareness affected by being on board one of the platforms? These questions were investigated in the first study described below.

In the AMUST concept, the co-pilot interacts with the UAV. In future operation concepts, e.g., Network Enabling Capability, 'sensors' and 'shooters' are allocated more dynamically. In a second study, briefly described here as well, we investigated the possible benefits for the pilot in having available a UAV image in the cockpit when performing a simulated Close Air Support mission.

6.7.2.2.1 Multi-Platform Situation Awareness

De Vries and Jansen [6] conducted a simulator experiment to assess the situation awareness of a co-pilot who controls a UAV platform while on-board a moving platform, in this case a helicopter. The co-pilot has to integrate the reference frame of the UAV with that of his own platform. Participants serving as UAV operators were seated behind a UAV console situated in a helicopter mock-up. The console displayed an electronic map of the geographical situation of a UAV, helicopter and a few other objects. For each simulator trial, the participants watched the movements on the electronic map display for a few minutes after which their situation awareness was assessed using an electronic questionnaire. Questions could be related to an earth-fixed coordinate system (compass) or be orientation-based (relative). Questions addressed the helicopter, UAV, or a formation of tanks. The main dependent measurement was angular error, i.e., the difference between the direction indicated by the participant and the real direction. The main experimental manipulation was a representation of the outside scenery on a 180-degree cylinder-shaped screen. Indicators were presented to show the movements of the helicopter. The simulation was presented in two work environments for the UAV operator: in a ground control station setting and on-board a moving platform (helicopter).

The most important results from this study are presented in Figure 6-37. The Figure depicts for the six different question types the angular error in indicating the location of an object. For most of the question types, the right bars (corresponding with the condition of the UAV operator aboard the helicopter) are much higher than the left bars (corresponding with the situation of an operator in a ground control station).

This indicates that in general terms, performance is worse in conditions in which we simulated that the UAV operator is aboard one of the moving platforms (except when questions were asked with respect to the platform itself). Note that the task itself is not different for the stationary and moving operator: just look at the electronic map and build up a spatial awareness as good as possible. The results revealed that it is indeed more problematic to maintain a multi-platform situation awareness while being aboard a moving platform. Apparently, when the operator is in a situation where one spatial frame of reference is dominant (i.e., while aboard a helicopter), it is very hard to process spatial information from other perspectives.

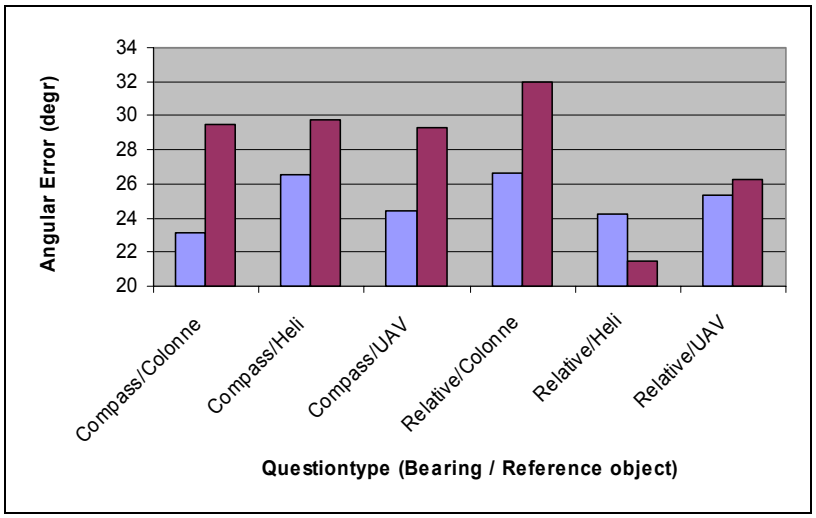


Figure 6-37: Angular Error in Indicating an Object Location for the Six Question Types. The left (blue) bars refer to the situation of an operator in a ground control station; the right (red) bars to the situation of an operator aboard a simulated aircraft.

6.7.2.2.2 Using Real-Time UAV Images while Conducting a Close Air Support Mission

The above research has shown that performance drops when multiple reference frames need to be integrated while performing a spatial situation awareness task. In a second experiment [7], a more critical situation was investigated in which misinterpretation of spatial information directly resulted in failure of a simulated Close Air Support mission. A situation was investigated in which a pilot used a UAV image, presented in the cockpit. As the UAV camera has a viewpoint that differs from that of the pilot, the pilot's interpretation of the spatial layout of the scenery may be prone to error (e.g., if the UAV flies in the opposite direction, the object on the left in the sensor image is actually on the right in the pilot's perspective). The aim of the research was to minimize the chance of such errors by rotating the UAV sensor image such that its orientation is always aligned with the pilot's spatial frame of reference. In this study, four military pilots performed several Close Air Support missions with six different display configurations. The orientation of the electronic map was either North-Up or Heading-Up. The UAV sensor image was either absent, present, but non-aligned (i.e., unadjusted image orientation: the image is presented as seen from the UAV viewpoint), or present and aligned with the orientation of the electronic map (adjusted image orientation: the image was presented as if the sensor was placed on the helicopter).

The pilots reported that they generally preferred the Aligned UAV sensor image in combination with a Heading-Up map. This preference was reflected in their performance, depicted in Figure 6-38: targets were identified twice as fast. Strikingly, flying performance was also better when an aligned UAV image was used.

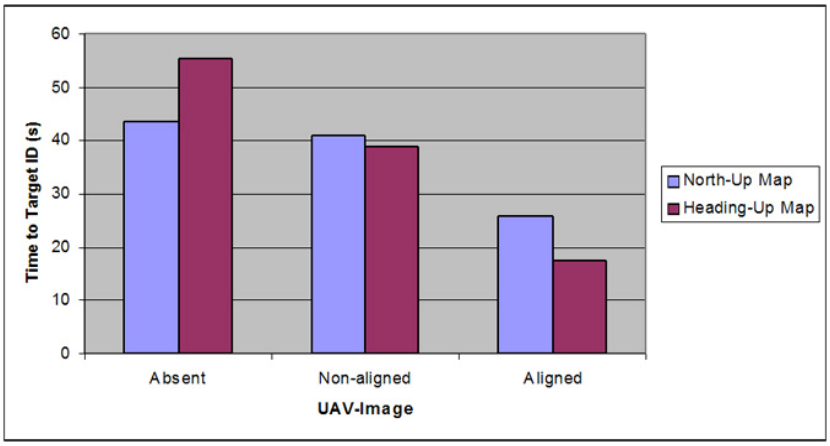


Figure 6-38: Time to Target Identifications for the Six Display Configurations.

6.7.2.2.3 Conclusions

In comparing the (simulated) situations of a stationary UAV operator and a moving UAV operator it appeared that it is very hard to process spatial information from multiple perspectives simultaneously when one perspective is dominant (here, when being aboard a helicopter). In experiments on presenting UAV images in the cockpit while performing a Close Air Support mission it was investigated whether the interpretation of (the non-dominant) UAV image information was facilitated by aligning that perspective with the dominant perspective of the helicopter. Based on the results and pilot's reports it was concluded that the availability of a UAV sensor image in the cockpit of an aircraft only improves mission performance when its orientation is aligned with the aircraft.

6.7.2.3 References for UMV Operation from Moving Platforms

[1] Walter, B.E., Knutzon, J.S., Sannier, A.V. and Oliver, J.H. (2004). Virtual UAV Ground Control Station. Proceedings of the American Institute of Aeronautics and Astronautics: 3rd "Unmanned Unlimited" Technical Conference, Workshop and Exhibit. AIAA 2004-6320.

[2] McCauley, M.E. and Matsangas, P. (2004). "Human Systems Integration and Automation Issue in Small Unmanned Aerial Vehicles." Naval Postgraduate School.

[3] Reason, J. and Brand, J.J. (1975). Motion Sickness. London: Academic Press.

[4] Sharkey, T.J., McCauley, M.E., Schwirzke, M.F.J., Casper, P. and Hennessy, R.T. (1995). The Effects of Whole Body Motion, Head Mounted Display, and Hand Control Device on Tracking Performance. Warren, MI: U.S. Army Tank and Automotive Command.

[5] Fayaud, G.R. (2001). The airborne manned unmanned system. Unmanned Systems, 19(4), 16-21.

[6] de Vries, S.C. and Jansen, C. (2002). Situational awareness of UAV operators onboard of moving platforms. Proceedings of the International Conference on Human-Computer Interaction in Aeronautics, HCI-Aero 2002 Cambridge, MA, 23-25 October 2002 (S. Chatty, J. Hansman, and G. Boy, Eds). AAAI Press: Menlo Park, CA. pp. 144-147.

[7] Jansen, C., de Vries, S.C. and Duistermaat, M. (2005). Presenting images from an unmanned aerial vehicle in an attack helicopter cockpit. Report DV3 2005-A16. Soesterberg, The Netherlands: TNO Defence, Security and Safety.

[8] North Atlantic Treaty Organisation, (1990). Standardization Agreement: Extreme Climatic Conditions and Derived Conditions for Use in Defining Design/Test Criteria for NATO Forces Materiel, MAS/048-MMS/2895.

6.7.3 Other Issues: Latency, Trust, Bandwidth

As well as the factors mentioned above, there are also a number of other considerations that have human factors implications for a UMV system. This section of the report summarises a few of the factors that are more likely to be encountered.

6.7.3.1 Latency

In this instance, latency refers to the time delay between the UMV operator making a control input and the feedback received by the operator from the system to confirm that the control input has had an effect. Such latencies are particularly noticeable, in the UMV context, when pointing sensors and issuing commands.

Early research suggested that as latencies increased, operator performance decreased in a linear fashion [3], however later research showed this to be oversimplifying the issue. For example, Asbery [1] varied the latency in feedback between 28 ms and 2000 ms for a target tracking task, and found that there was a sharp initial decrease in performance with latencies up to 100 ms with the detriment on performance diminishing thereafter until 1000 ms where it levelled out at a consistently poor level. This suggests that the claim of a linear relationship to be inaccurate. Asbery also found that the operators employed strategies to cope with a high degree of delay between action and response. When latencies increased, the participants seemed to lose sense of the orientation of the camera system. To overcome this, some participants delayed their inputs until the camera ‘caught up’ with their previous action. In general, trajectories became oscillatory in nature because of the overshoot generated by the high latency.

It has also been shown that latency may have very precise ‘effect boundaries’ where a certain range of latencies will have the largest effect on performance at a task. For example, studies have shown that for a simple control task (e.g., man-in-the loop UMV control) there is no noticeable effect of latency between 0 and 500 ms, but a sharp decrease in operator performance between 500 ms and 1000 ms. Similar patterns of results have been found in a trials investigating pilot’s landing ability using synthetic visual aids. In general, the findings from research in this area support the view that the impact a particular level of latency will have on a task is linked to the degree of control required to complete the task. A relatively low latency will have a much more pronounced effect on a task involving the operator making many minor adjustments than a task involving only general large-scale control movements.

Research on latency in virtual reality systems [2] found that latencies in update following a head movement had a more profound effect on the visual system than latencies following hand movement. It was suggested by the researchers that because hand movement is a feed-forward process, there is relatively little effect due to latency, as the control is not based on feedback from the visual system. On tasks that require visual feedback to monitor the results of actions, latencies have a much more noticeable effect.

To conclude, latencies do not correlate linearly with performance. They have specific ‘effect boundaries’ that will depend on the type of task undertaken. Latencies may have a less pronounced effect if the task requires

only feed-forward control and not visual feedback. Tasks requiring fine motor control are in general more severely impacted by shorter latency periods. Specific guidance on the likely acceptable level of latency for a given type of task is given in various standards. For example, UK Defence Standard 00-25 [4] lists the following recommendations for acceptable response times associated with human-computer interaction:

- <50 ms From movement of pointing device to movement of cursor or marker.
- 100 ms Pressing a key to character being displayed.
- 200 ms Selection of displayed object (field, button, menu option) to object appearing as selected; Selection of menu header to menu being displayed; Selection of scroll button to completion of scroll of one line of text.
- 1-2 s Completion of user input to display of error indication.
- 2 s Request for next page of information to completion of one page change; Completion of user input to completion of simple process; Completion of display manipulation request to completion of display change (e.g., open a window; zoom).

6.7.3.2 Trust

The UMV operator is required to make accurate and timely decisions. As they are physically removed from the vehicle they are controlling, they are reliant on the information provided to them from external sources, most likely from the UMV itself, and their experience of similar situations to make these decisions.

Trust relates to the operators' willingness to rely on the information presented. Clearly the operators must trust that the information provided is accurate if they are to use it effectively in the decision making process, rather than trying to 'second-guess' the situation and relying solely on previous experience. In addition, in UMV systems that rely on the operator acting as a 'system supervisor', the operator must trust that the automation in the system is carrying out the tasks associated to it in order for this mode of operation to work effectively. Without this level of trust, the operators' workload would increase significantly as they attempt to monitor the system too closely.

A discussion on decision making is outside the scope of this report, although the reader is encouraged to read one of the many texts available on human decision making. Instead this report will focus on issues and research directly relating to trust.

There are a number of theories of trust. For example, Muir [8] noted that the study of trust was under researched in the engineering psychology literature. She suggested that while the research may be lacking for trust in Human Machine Interaction (HMI), it would be possible to extend existing theories of interpersonal trust. Muir built on the theories of Rempel, Holmes and Zanna [9] to propose a three factor theory of how trust develops. Early on in the operator-machine relationship, the *predictability* of the system's decision is important in building trust. A system that makes unpredictable decisions will not be trusted. Once the system is perceived as predictable, an operator will trust a system more if it is deemed *dependable*. Dependability evolves over time as the system operator makes generalisations about the specific actions of the system to a broader set of attributes of the system. Dependability will be particularly strong if the machine is able to perform accurately under risky situations. The final stage involved the development of *faith*, the belief that the system will be dependable in the future. During this stage of the 'relationship' between the operator and the UMV, transparency in the UMVs decisions will be important in maintaining trust.

This, and other models of trust applicable to UMV systems, tends to neglect the characteristics of the operator in favor of aspects of the system. For example Kelley et al.'s model [10] does take the operator into account to a limited degree. In this model, trust is gained and influenced by the *competence* of the system (as summarised by earlier models, described above), *understanding* and *self-confidence*. Self-confidence contains factors such as skills, experience and faith which directly relate to the operator of the system. Similarly understanding contains factors that arise as a result of the operator and UMV in combination; predictability and familiarity for example. This model was originally constructed to understand trust within the Air Traffic Management (ATM) environment, although it can be applied to many areas of human machine interface. The model also examines trust at a broader level, fitting it into a model of system acceptability alongside factors of teamwork and situation awareness (Figure 6-39).

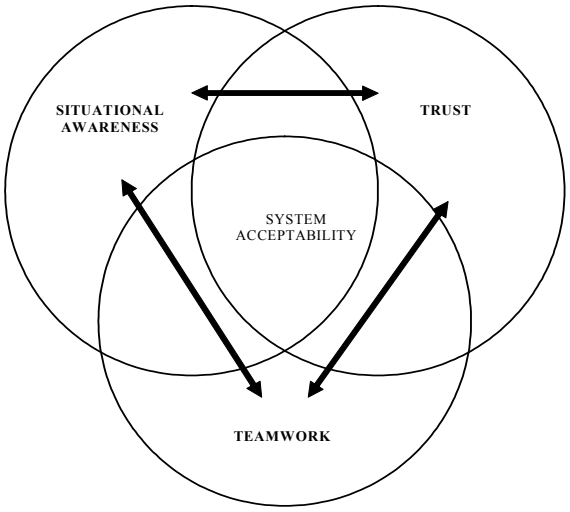


Figure 6-39: Trust in ATM Context (Kelley et al., [10]).

Most models are lacking, as they generally ignore the properties of the operator. Others include the operator, but only when factors such as confidence and training have been explored [5]. Additional attributes of the operator may also have an impact such as individual differences in overall propensity to trust.

6.7.3.3 Bandwidth and Update Rate

The bandwidth of the communication link between the operator and the UMV and the update rate of the information that is provided are important considerations in UMV system design. In order to control a system effectively, the operator requires the right amount of information to be provided at the right time.

Bandwidth limits the size of the information passed down the communications link, whilst update rate refers to how often the information is refreshed. Both are interlinked; low bandwidth links may impose a limit on the speed that information can be transmitted. Bandwidth may be restricted by technological limitations or by operational restraints, for example because of a need to operate the system covertly.

Work completed by Haduch [7] suggests that the quality of feedback provided to an operator will effect how the UMV is operated. Of all the senses, visual feedback is most often favored by system designers for feedback, possibly because it is perceived as being the most informative. Haduch suggests that it may be

possible to utilise a form of cross-modal communication; using audio communication rather than visual to represent speed for example. Haduch also suggests ways of reducing the communications bandwidth required to operate a UMV. A simple example of this is reproducing vehicle generated sounds at the control station. This way, the actual data of the sound of the wheels spinning, for example, need not be sent (as this could place heavy demands on the bandwidth available), but merely a message indicating to the control station to make a sound of wheels spinning. Such a system may not be applicable to all situations, but will lessen the need for large communication bandwidths in many situations.

Bandwidth requirements will also vary depending on the operators' role. A supervisory role with intervention from the operator only in emergencies, will cope much better with smaller bandwidth communications links than a system that requires full operator in the loop control. This implies greater implementation of automation into UMVs which in turn has human factors considerations (e.g., see above), but will lessen bandwidth requirements.

Any bandwidth reduction would initially require investigation into what the key information needed for the task was. This information is relatively lacking in the literature, although this is most probably because the key information will the highly task dependant [6].

6.7.3.4 References for Other Issues

[1] Asbery, R. (1997). The Design, Development and Evaluation of an Active Stereoscopic Telepresence System. PhD Thesis. Guildford: University of Surrey.

[2] Kawara, T., Masao, O. and Yoshizawa, T. (1996). Effects on visual functions during tasks of object handling in virtual environment with an head mounted display. Ergonomics, Vol. 39, No. 11, 1370-1380.

[3] Warrick, M.J. (1949). Effect of transmission-type control lags on tracking accuracy. (Technical Report 5916). Dayton, Ohio: Wright-Patterson Air Force Base, Air Force, USAF Air Material Command.

[4] Defence Standard 00-25, Human Factors for Designers of Systems (Part 19, page 163), Issue 1.0. Defence Procurement Agency.

[5] Adams, B., Bruyn, L., Houde, S. and Angelopoulos, P. (2003). Trust in Automated Systems Literature Review. Humansystems Inc. on behalf of Defence Research and Development Canada.

[6] Willis, A.R., Beagley, N., Davies, I. and Stringer, N. (1999). Remotely Piloted Vehicles: A Review of Human Factors Research and Recommendations for Research. DERA/CHS/PPD/CR990013, Defence Evaluation and Research Agency.

[7] Haduch, T.W. (1993). Operator interaction with Unmanned Ground Vehicles (UGV) (with limited data transmission), In: Smith, M.J. and Salvendy, G. (Editors), Human-Computer Interaction: Applications and Case Studies, Amsterdam: Elsevier.

[8] Muir, B.M. (1994). Trust in automation: Part 1. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics, 37 (11), pp. 1905-1922.

[9] Rempel, J.K., Holmes, J.G. and Zanna, M.P. (1985). Trust in close relationships. Journal of Personality and Social Psychology, 49, pp. 95-112.

[10] Kelly, C., Boardman, M., Goillau, P. and Jeannot, E. (2001). Principles and Guidelines for the Development of Trust in Future ATM Systems: A Literature Review, European Organisation for the Safety of Air Navigation, pp. 1-48.

6.8 SUMMARY

This chapter described many candidate control/display technologies that are potentially applicable to UMV operator interface design and detailed other important design considerations. However, since UMVs encompass a very broad range of vehicle types, capabilities, mission contexts, and environmental constraints/ affordances, it is important that any UMV operator interface design follow a multi-disciplinary user-centered design process. The goal of user-centered design is to ensure the final design meets the users' needs and expectations. The process of requirements definition (user profiles, work flow, task analysis, and information architecture) and repeated interface design development and iteration (through multiple usability assessments and formal evaluations) will increase the likelihood of obtaining a truly functional and easy-to-use interface.

Other key conclusions can be drawn from this chapter, besides the importance of user-centered design. First, the potential importance of mission-specific multi-modal interfaces to UMVs is highlighted. Since UMV operators are currently limited to a reduced stream of sensory feedback delivered almost exclusively through the visual channel, there is reason to believe that situation awareness and performance may be improved through increased multi-sensory stimulation. These improvements might stem from an increase in the operator's sense of 'presence' in the remote environment, from increased information throughput afforded by effective use of multi-sensory stimulation, and/or a more intuitive presentation/control of information, and thus improved performance over conventional visual interface practices. Technologies such as spatialized audio, haptic/tactile stimulation, and speech recognition systems appear especially relevant to multi-UMV operations.

Additionally, as UMVs become more automated/autonomous and single operator control of multiple UMVs is mandated, the importance of the human-automation interface becomes paramount. As mentioned above and in Chapter 7, there is a wealth of potentially negative effects associated with human-automation systems that have significant implications for the operator interface design. Automation must be designed to augment, not hinder, human capabilities. Operator interfaces must provide rapid visibility into the current status and future plans of automation for shared human-automation situation awareness. Additionally, intelligent decision support interfaces will need to be designed such as to allow independent operator assessment of the situation as well as the rationale for any automated classifications/recommendations.

Finally, the chapter clearly indicates several areas in need of further research. These areas include the relative costs/benefits of the various control/display technologies identified for particular classes of UMV applications (air, ground, sea), a better understanding regarding the application of multi-modal interface technology for UAV operator interfaces, the issues surrounding human supervisory control, and the design of effective decision support interfaces for enabled multi-UMV control.