ARI Research Note 97-16

Enhanced Virtual Presence for Immersive Visualization of Complex Situations for Mission Rehearsal

Paul Cutt
Xtensory, Inc.

Research and Advanced Concepts Office
Michael Drillings, Chief

June 1997

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Xtensory, Inc.

Technical review by

George Lawton

NOTICES

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1997, June	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) September 1986-August 1987
4. TITLE AND SUBTITLE Enhanced Virtual Presence for Immersive Visualization of Complex Situations for Mission Rehearsal		5a. CONTRACT OR GRANT NUMBER DASW01-96-C-0046 5b. PROGRAM ELEMENT NUMBER 0601102A
6. AUTHOR(S) Paul Cutt		5c. PROJECT NUMBER B74F 5d. TASK NUMBER 6901 5e. WORK UNIT NUMBER C02
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Xtensory Inc. 140 Sunridge Drive Scotts Valley, CA 95066		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 97-16

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: N/A

14. ABSTRACT *(Maximum 200 words)*:

This final report describes the key requirements of an internet-based system that provides an immersive environment for mission rehearsal. It shows how off-the-shelf hardware and software can be used to meet those requirements.

A layering implementation technique is used with the common hardware and software to provide economies of scale both in the use of new hardware/software and functionality.

Finally, specific hardware/software is described where necessary to provide support for mission rehearsal. In particular, new inventions such as the Internet Appliance and Remote Access Controller are described in detail together with their implementation.

15. SUBJECT TERMS

Virtual reality   Visualization   Simulator   Immersive   Mission rehearsal
Training   Virtual environment

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	67	

Table of Contents

Preceding Page Blank

List of Figures

Summary

This final report describes the key requirements of an internet-based system which provides an immersive environment for mission rehearsal. It shows how off-the-shelf hardware and software can be used to meet those requirements.

A layering implementation technique is used with the common hardware and software to provide economies of scale both in the use of new hardware/software and functionality.

Finally, specific hardware/software is described where necessary to provide support for mission rehearsal. In particular, new inventions such as the Internet Appliance and Remote Access Controller are described in detail together with their implementation.

Overview

Requirements

The work for this project was to develop the requirements for a system that would provide the capability for "Enhanced Virtual Presence for Immersive Visualization of Complex Situations for Mission Rehearsal." One key criterion was that the system should be composed of off-the-shelf hardware and software to the maximum extent possible.

After examining field reports provided by the Army (see Appendix A - Unit Performance Summary) it became apparent that a key problem area for mission rehearsal was that of coordination. In many instances plans and schedules were developed in isolation by different battlefield groups, that when required to operate in synchronization, failed to mesh or interfered with each other. In many cases, resources assumed to be present were not there at the times and places required. This area requires the integration of schedules, plans, and data.

The second area that we examined was that of presentation of information. In a battlefield context, information overload becomes a key constraint. The issue was how to hide the information, yet make it available to the user when required. We proposed the use of various virtual reality techniques that can be used to "hide" information, yet be instantly recallable. Related is the problem of highlighting information whose status has changed to critical.

The third area is that of communication. A mission may bring together different groups from different services such as Air Force or Navy. Any mission rehearsal system needs to have facilities to enhance communication and bring a sense of "community." Related to this is ease of use. The system must not be so hard to use that it gets in the way of communication. In fact, since this system may be used in battlefield conditions, it would be critical that its use is in a sense intuitive so that users not trained in its use can be proficient quickly.

Another aspect of communication is that of capturing information as it flows directly from the battlefield. The more recent and accurate the battlefield data, the more likely is success.

Also, we do not have time to examine all the implications in this report, but we believe that a fruitful avenue to explore is how such a system can be used to enhance the sense of community in planning for and rehearsing a mission. The more a sense of "belonging" and "sharing" the system promotes, the better the chance of success. In many instances, especially in complex situations, different teams are brought together and may work together for the first time. The quicker a sense of "community" or working together on a common mission is developed, the smoother the final operation. We have just started exploring use of avatars as a means to enhance the sense of community.

We also considered the "rehearsal" aspect of the system. We wanted a system that is easily accessible by the different groups in the mission rehearsal. This means that any and all control centers must be easily moved from point to point with no single point of failure.

We also wanted the ability to capture "virtual missions" for archiving in a database. With the ever-changing political landscape, it would be nice to start keeping in a database future possible sites for missions. In many cases, e.g. Iraq, western firms were instrumental in planning and designing many of the buildings, structures, roads and other infrastructures. This information, when gathered, could be used to plan future missions. To avoid information overload, all that is needed to be kept is links to the actual information, not necessarily the information itself. The greater the amount of information about the mission site, the better the chance of success.

Another aspect of the rehearsal that we were interested in was the frequency and ease of rehearsal. We took as our starting examples Israel, Switzerland and Singapore. These are small states, and the need to maximize their military resources is essential. A key aspect was that these countries had a large reservist population that could be mobilized quickly. What if the system proposed here could be used to mobilize the groups involved in a mission as efficiently if not better? The system would be more efficient if it could be used in any location, e.g. at home, yet be able to provide the ability to rehearse at least part of the mission.

Feedback from the mission participants must be accounted for. Post mortem analysis is an important feature of any project, and is especially useful in a mission rehearsal context. It is from evaluation of past actions that we, and others, learn. It is important that feedback is integrated into the system providing for continued learning, communication, coordination and information.

System

We proposed a system that would be based on internet technology and uses virtual reality to tie the different components together and provide the "enhanced presence." The Engagement Response System consists of servers, client enhancements and network appliances using internet technology:

- An *Engagement Response Server.* The servers (there may be more than one) serve as control centers to collate, display and distribute the mission data to the clients. Since this is based on the internet, the clients would be a system with an enhanced network browser.

- *Enhancements to the internet network browser* such as Netscape Navigator or Microsoft Explorer. Specific enhancements would be those to VRML 2.0, 3D sound, QuickTime VR, terrain/maps, spherical coordinates, etc.

  One key enhancement we have identified as critical are the translators. These translators are used to convert schedules, plans and other data from different groups involved in the mission to a common format for the mission.

  Another enhancement is the gathering and coordination of plans and schedules. In this we are guided by the architecture/engineering/construction (AEC) industry. For actual testing of the system, other commercial uses may be more relevant, such as fire fighting, police SWAT teams, etc. However at this stage, the AEC industry is important because part of any mission rehearsal is the support infrastructure. For example, in Desert Storm, a key part of that mission was the logistics of support, such as roads, bridges, equipment that had to be brought together in a timely manner.

- *The Internet Appliance*, a term we developed to describe a tool to gather mission data for the servers. The Internet Appliance is a very small PC-based system which can be loaded dynamically with the programs to collect the mission data of interest.

  One example of its use is to be loaded with GPS and QuickTime VR collection programs. When so loaded, the Internet Appliance together with a digital camera can be sent into the field to the mission site. There the user can capture visual scenes of the mission site. The user can then, or at a later time, dial into the internet and transfer the data into the mission database.

  If the Internet Appliance is loaded with a different set of programs it may be used to acquire "ground truth" information where ground truth is the description of a user in the virtual environment encompassing items such as velocity, vehicle type, armament, radar signature and others.

Operations

The Engagement Response Server and clients can run on any system which supports the internet, which effectively means any computer, either workstation or PC based, Unix or Windows 95/NT. In fact, both the server and client can run on the same computer. Where they run is dynamic so a computer can be a client at one instance and a server the next.

Design Philosophy

We would like to emphasize the design philosophy. The design philosophy is to layer enhancements, both hardware and software, onto existing off-the-shelf products. In this, we follow the design philosophy used by another of our products, SyncLink™. In SyncLink, we layered a common virtual reality device API onto most virtual reality devices. This provided a common interface for virtual reality devices which then could be used in existing applications while the underlying virtual reality devices could be switched or altered without affecting the application. The use of common translation and transformation routines made transparent which virtual reality device is used.

The success of this approach is shown by the diverse applications in which our customers are using SyncLink. For example, Lockheed is using SyncLink in military applications, the Mayo Clinic in medicine, the US Navy in undersea telerobotics, Disney in entertainment, and Deneb in robotics.

Using a similar approach, the Engagement Response System can take advantage of the ever increasing functions of the internet. If the layering is done carefully, as the internet increases its capabilities, so does the Engagement Response System.

The advantages of careful layering cannot be over emphasized. For example, more and more databases are becoming accessible over the internet. So, data from Oracle, Microsoft Access can be accessed. It is highly likely that data for a mission will be stored in an industry-type database such as Oracle or spreadsheets such as Excel. If so, the access by the Engagement Response System for mission rehearsal becomes easy.

Another final example is the convergence of the PC and TV to provide internet-capable TV. We now have large numbers of persons trained to use the internet which lessens the training for persons to use the Engagement Response System, but even more importantly, we have a means for mission rehearsal from individuals wherever they may be, as long as they have an internet-capable TV and access to the internet.

Our vision is to create the Engagement Response System as an add-on similar to a "plug-in" to a standard internet server and browser.

There are many advantages to using such an approach. Xtensory has been designing such a system for the construction industry which has similar coordination and scheduling problems as the Army. Some of these advantages are :

- Integration of most data input sources such as video, fax, voice and graphics input comes at no cost

- Whiteboarding is available for conferencing

- Scheduling is immediate and available to all. This is NOT standard internet technology as yet but we have designed the ability to take schedules directly from a variety of sources such as Microsoft Project, Microsoft Word and scanned images, and coordinate the presentation so that navigation through a complicated set of schedules is available to all authorized parties and is updated constantly from a connected data source.

- Data can be taken directly from databases such as Oracle, Paradox, Access and other standard databases so the commanders and others get immediate data.

- Through appropriate design of the engagement browser we can present different views of the same data to different users.

The above features can be used for fire support plans, air defense planning and coordination and support between company teams.

We studied the commercial off-the-shelf hardware and software of the internet, and of the web in particular. This included literature studies and discussions with manufacturers. The studies also included actual use of the web. Using the software and hardware defined, we set ourselves up as a web host on a web server and worked with the following:

- Web page design
- Databases
- Forms entry
- Email response
- Scripts (CGI, Java, Perl)
- Password/firewalls
- Remote access
- Audio
- Video
- Videoconferencing
- Whiteboards
- Animation.

This was done in using primarily Windows NT with some work on Unix.

We acquired or purchased internet/web tools for this exploration:

- HTML authoring software (SoftQuad HotMetalPro, Corel)
- Web server (Netscape)
- Web browser (Netscape)
- Databases (Microsoft Access, Oracle)
- VRML authoring software
- VRML browser
- Audio and video software (Connectix, CU-SeeMe)
- Whiteboard software (Connectix)
- Video cameras (Connectix)
- Sound boards (Gravis Ultrasound Max)
- Animation.

As we continue identifying commercial off-the-shelf hardware and software our list is likely to change.

Our web server, www.xtensory.com, consists of a Pentium PC running Windows NT and Netscape Server. It resides in our lab and is connected to the internet via a 56k frame relay phone line with a router and CSU/DSU.

Through careful design it will be possible to embed the Engagement Response System within the server and browser in such a way that additional features are added to the standard server and browser technology such as reconnaissance route overlays, terrain maps and scheme of maneuvers.

The Engagement Response Server

Definition of Engagement Response Server

Although the Unit Performance Summary document is indicative of the requirements presented to us, it is certainly not believed to be the only set of requirements. Other kinds of events or practices will impose other requirements on a Engagement Response Server. Various demands can be placed on the Engagement Response Server, including

- Purpose
  - Mission rehearsal simulation
  - Planning for mission rehearsal simulation
  - Evaluation of mission rehearsal simulation
  - Playback of mission rehearsal simulation
  - Command center in real time for real-world simulated combat exercises, or even real actions
- Key User
  - High level (e.g. battalion commander)
  - Low level (e.g. infantry soldier)
  - Intermediate
  - Mixture of various levels of users
  - Simultaneously having various levels of users using the system
- Number of participants
  - Single
  - Multiple
- Location of multiple participants
  - Co-located in same place
  - Distributed in geographically
- Model Sources
  - Virtual environment and objects programmed in (e.g. CAD models)
  - Virtual environment and objects from real world (e.g. photographs, videos, radar, etc.)

- Mixture of virtual and real environment
    - Getting inputs from real world
    - Non-traditional models and representations
- Model Relationships
    - Internally generated inputs: Virtual environment and objects controlled only by user actions, or by actions within the virtual environment
    - Externally generated inputs: Virtual environment and objects controlled by information and sensors from real world (e.g. GPS sensors on real-world vehicles send position information to virtual environment)
- Control
    - Participant(s) in control. The Engagement Response Server responds to their actions.
    - Alternatively, the Engagement Response Server can be running as a simulator engine with much of the activity in the virtual environment programmed in and the user(s) responding to it.
    - Mixture
    - Opposition forces
        - Control of foe is by foe
        - Control of foe is by computer
        - Mixture
- Behind-the-scenes commander (e.g. superuser, Engagement Response Server manager, God) can dynamically change parameters in real time
- Interaction
    - Between user(s) and computer
    - Between multiple users (communication)
    - Discussion groups.

Visual Presentation Requirements

The purpose of visual presentation of data would be to provide "engagement awareness" by aiding attention focus. In particular, to identify the key elements of the information as it comes in so that it may be assimilated and assessed. The view of such a system would be to provide an "engagement awareness" to the commander so that the engagement is "mapped to body awareness." The commander is part of the engagement. Key elements of the engagement such as disposition and capabilities of the enemy are brought automatically to his attention when warranted and appropriate and relevant cues provided through sifting the large stream of information fed on the battlefield and coalescing of the information

We would like to identify some of the requirements. The Engagement Response Server (ERS) requires the following elements:

- A means for grouping components of the data into categories. ERS needs to provide the ability to divide the set of possible interactions with the data into a number of smaller sets with disjoint different dimensions

- The ability for each category to raise or lower the level of "user awareness."

We assume this system sits on top of a virtual reality rendering engine which already provides the ability to provide terrain and physically accurate building maps from input terrain description (e.g. Department of Defense digital terrain and elevation data) and converts it to polygonal data. See Figure 1. Through the use of the virtual reality engine, the ability exists to move to a location in the battlefield, observe the activity there, and review. Our interest lies rather in the integration of the non-physical data such as radar threat envelopes or fields of fire with the physical data, and in particular, with the organization of such non-physical data so that it enhances the "engagement awareness" of the commander.

In fact, the ability to hide and mask temporarily undesired portions of the environments is as important as the ability to highlight portions of the environment. The system should also map situations to body motions. The next steps are defined further:

- Identify a "situation." We define an engagement to consist of many time- and spatially-varying situations.
- Obtain real data to set up a database for a "situation"
- Obtain a virtual reality engine to provide terrain and building maps
- Develop demo categories and groups
- Set up auctions
- Get feedback (psychological/behavioral aspects)
- Provide test system components to test assumptions.

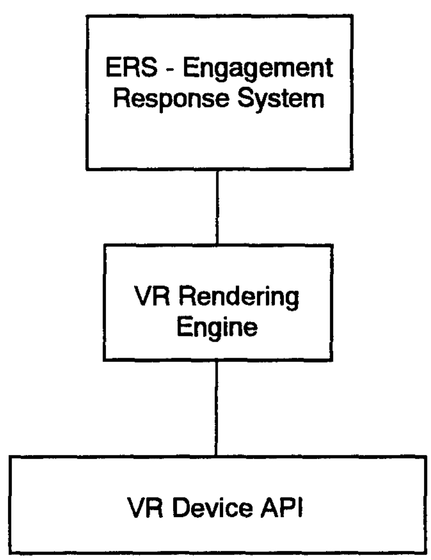


Figure 1: Block Diagram of Virtual Reality System Incorporating the Situation Response System

Data Grouping

Components

To enable the grouping of data, virtual reality provides some unique techniques. These techniques which have been used by others in the past and which will be exploited here are:

- Virtual Menus
- Portals
- Search and Discrimination
- Miniature Worlds within Worlds
- N-Maps
- Multiple Viewpoints.

Events in an engagement are not isolated; they are related to other events in time and space. A set of events related either temporally or spatially form a situation. These situations together form the engagement.

There are two main areas that need to be addressed. The first is the organization of these events and situations so that they make logical sense to the commander and that they be "enabled" to make their presence known to the commander where appropriate. We term this "situation awareness." The other area is the area of communication between the commander and his support staff to ensure that "situation awareness" is conveyed and that consensus is obtained.

Components of any situation must be able to be both:

- *Time sequenced.* In time sequencing, the components can be sequenced in time. For example, menus can be presented sequentially in time, one after another.

- *Spatially Sequenced.* In spatial sequencing, the components are positioned in 3D space. The spatial positioning provides the context. Size plays a part in any space, and with it the requirement for changing sizes, accomplished by *zooming* from a large area to a small area and vice versa.

Component Description

- Using *virtual menus* the user can define a set of menus that are 3-dimensional menus, which can be positioned in 3D space, so that he is surrounded by these menus. He can interact with these menus, and by appropriate grouping of these menus in the space, he can set up a command center that literally floats around him. Work in the area of virtual menus has shown that they can be an effective means of interaction with the virtual environment (Jacoby, 1993). However, some amount of flexibility is needed to provide access to the menus without restricting view within the virtual environment.

- *Portals* provide a means of gathering various components into logically organized groups. So within a portal, you might find previously defined scenarios which can be plugged into the current situation.

- Providing a *search and discrimination* function is essential since the costs of searching and discriminating increases exponentially with the number of objects on the battlefield. To provide such a function, components must be capable of being organized into groups which are then are combined to provide entities. The search and discrimination function we choose to concentrate on at present is an auction type function. Using an auction mechanism, these entities organize among themselves and bid for situation awareness. Other techniques that we know of have been used to provide search and discriminant functions are "fuzzy logic" techniques. But as far as we know, using market based mechanisms such as "auctions" to organize the entities is still relatively new. The key features of this technique is its familiarity to many persons and so as such requires very little training for a person to be effective. This is very important as during the situations under which ERS is to be used, it is highly likely that a person untrained in the system might have to use it.

- *Miniature Worlds within Worlds* provides a way to visualize n-dimensional information in a 3D virtual environment. By embedding a 3D world within another 3D world, higher dimensions can be represented. The multiple worlds are visible, so the position of one virtual world with respect to another represents three of the variables. This can be readily visualized as nested 3D graphs (Feiner, 90). The choice of axes as well as the nesting order dramatically affects the surfaces represented, so means of organization have been developed to facilitate ease of use. In addition, virtual tools such as a dipstick, waterline, and magnifying box can be used to interact with the nested 3D graphs to find specific information.

- *N-Maps* are containers for presenting graphical output and capturing multivariate data. They can be used to present threat envelopes displayed for vehicles or radar envelopes. A box's coordinate system represents a transformation relative to its parent. See Figure 2. Each box is an instance of a class that may be associated with event handlers that allow it to register and react with different events. A box can map and unmap itself from the battlespace, and mapped boxes are displayed. A controller box is one that owns a child box and maps or unmaps it as it sees fit. This implements the schematic standing for a more complex object.

- *Multiple Viewpoints.* The ability to position multiple viewpoints at various positions of the battlefield so that the commander can move quickly to different locations and locate items of interest is basic. We assume the ability to "set book marks" in various parts of the battlefield.

Situation Awareness

Throughout the engagement important portions of the data will vary from moment to moment. What part of the data is important may change drastically so providing a means where the importance and the "awareness" of the data to the user can be quickly changed is of extreme importance. The mechanism used or the "awareness function" to group the data must provide for this capability. In this respect, an "auction" type mechanism whereby data can essentially "bid" for "awareness" has an intrinsic appeal. The auction mechanism also provides a means for the user to increase or decrease the importance of the data to him. Conceptually, the components form a set of sometimes competing and sometime cooperating mechanisms that attempt to make the commander aware of missing or forgotten important information. See Figure 3.

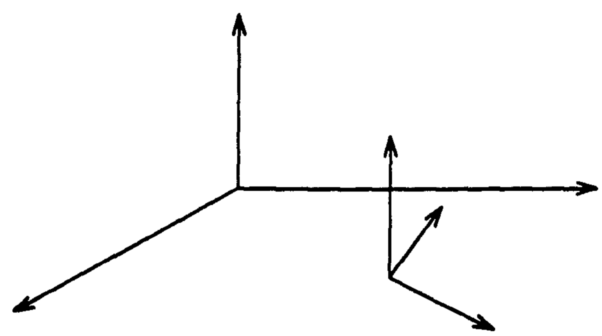
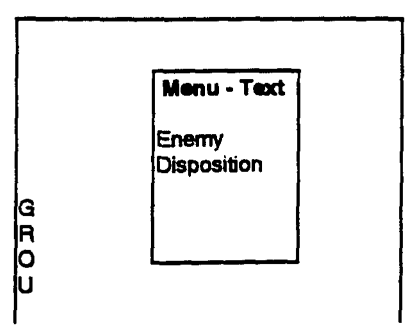


Figure 2: Box Hierarchy of N-Dimensional Space

PLEASE NOTE that although we have chosen to focus on one particular awareness function, in practice multiple functions could be used. So work being done using fuzzy logic, for example, would be relevant. Part of our interest was to find a simple, understandable, yet workable function which could be used with little or no training.

Awareness Function - The Auction Mechanism

Markets and how they function constitute the core of any economic system, whether it is highly decentralized (a "capitalistic" system) or highly centralized (a "planned" system). This is true for the decentralized economy because markets are the spontaneous institutions of exchange that use prices to guide resource allocation and human economic action.



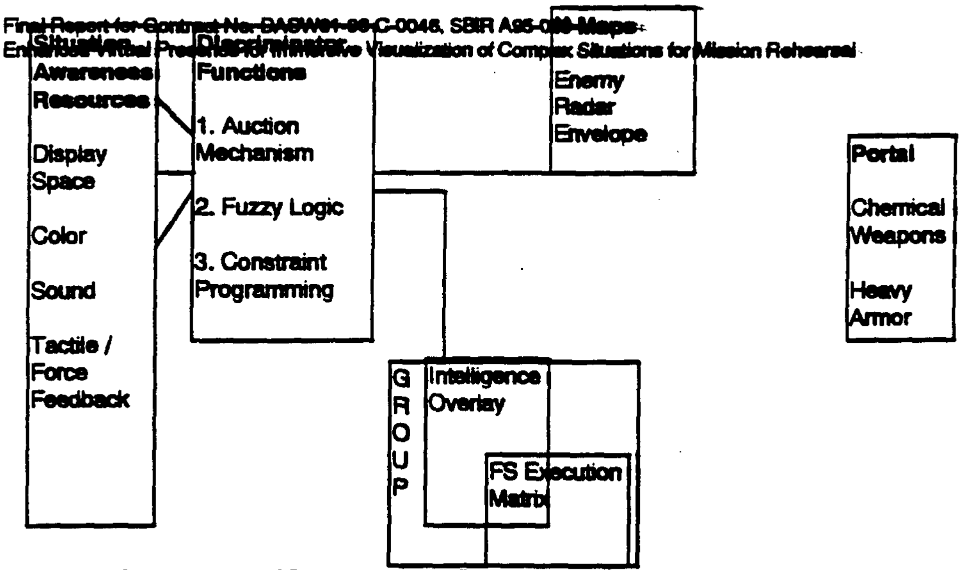


Figure 3: Schematic of Engagement Response System

For example, a group of components representing a data trend could be issued vouchers. One of these vouchers could be redeemed by the bearer as a cash substitute in the purchase of new display credits. A display credit enables the flashing of the group for example. Consequently vouchers are of value to future groups. Furthermore, since (as Hayek would say) the "circumstances of time and place" for the potential redemption of vouchers are different for different groups, there exists the preconditions for the active voucher market that will arise between groups as they bid for and exchange vouchers. Current groups with vouchers who are unlikely to be displaying again soon hold an asset worth less to themselves than to others who are more certain of their future or impending display plans. The resulting market establishes prices that are discounts from the redemption or "face" value of vouchers. Sellers who are unlikely to be able to redeem their vouchers prefer to sell them at a discount for display credits. Buyers who are reasonably sure of their display plans can save money by purchasing vouchers at a discount. Thus the welfare of every active buyer and seller increases via this market and this serves the true benefactor the user.

The previous paragraph illustrates a fundamental hypothesis (theorem): the ("competitive") market process yields welfare improving (and, under certain limiting ideal conditions, welfare maximizing) outcomes for situation awareness by dynamically bringing to the attention of the commander important data and trends without the commander having to specify in advance what is important. This is critical because in many situations the commander may not even be aware of the importance of a data item or trend. However please note the use of market mechanisms for this purpose is ONLY a hypothesis. The key here is that what is needed is that there is some way whereby the commander can be notified of the importance of some data/data trend even where he may not be aware of it's importance or the data's importance has only arisen during the course of the conflict.

Is the hypothesis "true", or at least very probably true? (Lakatos (1978) would correctly ask "Has it led to an empirically progressive research program?") I think it is "true", but how do I know this? Do you see what I see? A Marxist does not see what I see in the above interpretation of a market. The hypothesis remains to be satisfied by experiment as to its applicability to provide satisfactory "situation awareness" for an engagement response system.

In any actual experimentation, they are many variations based on real markets and real auctions. For example, setting of seller's minimum selling price, buyer's maximum buying price, defining trading periods and limits on sales. Thus demand, (supply, S), can be modeled as "renewed" in each trading period, a steady state flow, with no carry-over in unsatisfied demand (or unsold stock), from one period to the next. In the example above, imagine the vouchers being issued, followed by trading; the vouchers then expire, new vouchers are issued, traded and so on.

The provision of market preconditions that (1) "the circumstances of time and place" for each economic agent are dispersed and known only to that agent (as in the above voucher market) and (2) agents have a secure property right in the objects of trade enables the earning of private gains ("profits") from trade (a voucher was transferable and redeemable by any bearer). The fact that sellers and buyers are software programs makes no difference.

The 'profit' or the surplus earned by a buyer who buys for less than his willingness-to-pay, just as a seller's 'profit' is the surplus earned when an item is sold for more than the amount for which they are willing to sell can then be used to purchase display time. The willingness-to-sell, like willingness-to-buy, is determined by the immediate circumstances of each agent. Thus the earning of "profit" enables a group that has a high need to display itself to bid for and pay for the display it needs.

The auction mechanism is of course not the only awareness function that can be used. It may not even be a suitable one. The usefulness of such an auction mechanism needs to be determined by experimentation.

Communication - Consensus Between Users

Not all that is in a virtual environment is physical, as it is in the real world, and getting agreement among different users becomes an even more important issue. It is basic and important to be able to answer the question "Do you see what I see?" Replication and control are the two primary means by which we attempt to reduce error. However, the question "Do you see what I see?" contains three component questions, recognition of which helps to identify three different senses which may fail to be replicable:

(1) Do you observe what I observe? Since virtual reality has traditionally been confined to the presentation of physical data, e.g. buildings, the answer to this question has been trivially, "yes". Where non-physical data is concerned, there is in fact much more room for mistakes of observation.

(2) Do you interpret what we observe as I interpret it? Given that we both observe the same, or replicable data, do we put the same interpretation on these data? The interpretation of observations requires theory (either formal or informal), or at least an

empirical interpretation of the theory in the context that generated the data. Theory usually requires empirical interpretation either because (i) the theory is not developed directly in terms of what can be observed (e.g. the theory may assume risk aversion which is not directly observable), or (ii) the data were not collected for the purpose of testing, or estimating the parameters of a theory. Consequently, failure to replicate may be due to differences in interpretation which result from different meanings being ascribed to the theory. Thus two researchers may apply different transformations to raw field data (e.g. different adjustments for the effect of taxes), so that the results are not replicable because their theory interpretations differ.

(3) Do you conclude what I conclude from our interpretation? The conclusions reached in two different research studies may be different even though the data and their interpretation are the same. In economics this is most often due to different model specifications. This problem is inherent in non-experimental methodologies in which, at best, one usually can estimate only the parameters of a prespecified model and cannot credibly test one model or theory against another. An example is the question of whether the Phillips' curve constitutes a behavioral trade-off between the rates of inflation and unemployment, or represents an equilibrium association without causal significance.

Communication Requirements

A common thread running through the Unit Performance Summary (see Appendix A) is the importance of good communication, and the damaging results of poor communication. Communication, or exchange of information, between people is of utmost importance. It is one thing to have a single-user Engagement Response Server for some isolated purpose, but it is an entirely different situation when a Engagement Response Server has multiple participants. Communication between the participants is one of the key strengths of the Engagement Response Server. Any standard server can be provided with many different communication mechanisms designed with communication between participants as the fundamental requirement.

These mechanisms include:

- MUD type "worlds" using avatars where different users can meet
- Discussion groups
- Live broadcasting
- Video conferencing
- Telephony.

The communication between the participants can be in discussion groups, as are commonly found on the internet/web for effective interaction. These discussion groups can bridge the different stages of the virtual reality simulation, from planning to the combat simulation itself to post mortem analysis. The communication between the participants can also be by two-way audio and video broadcasting.

Information and Using the Engagement Response Server for Event Simulation

---

As we review what is required of a Engagement Response Server we are constantly reminded that we need to have enough information but can easily be overwhelmed by too much.

Representations that are not possible in the real world are possible in a virtual reality simulator and can be used to provide additional information. For example, areas that can be sensed by enemy's radar can be shown, alerting the user to stay away. If the virtual environment has inputs from a real battlefield, it may be possible to trace back indirect-fire and direct-fire weapons to their source.

Information about the simulated combat activities are varied and open-ended from the viewpoint of a Engagement Response Server designer. It is necessary as designers to provide the capabilities and tools necessary for the Army to succeed in having a useful Engagement Response Server. It is likely that there will be military requirements that are unknown to us, and it is just as likely that new requirements will be needed as soon as the Engagement Response Server starts being used. By providing an open system we can have a base system that can be enhanced as new requirements are generated.

Most of what we have been discussing has centered on information from the user's viewpoint, whether the user has been a battalion-level commander the low-level soldier. So, what does it mean to have that viewpoint? It typically means that the system is geared to one user, or toward one side in the battle. But the battle will have two sides, us and them. Certainly, in the simulated combat events, "we" want to use the Engagement Response Server for our benefit, but "they" can also benefit from use of the Engagement Response Server, too. The Engagement Response Server can be viewed as a two sided system, and each side can have access to the same tools as allowed by the Engagement Response Server manager. These tools, as noted above, included planning, the actual simulated combat activities themselves, review and summary, plus analysis and critiques. Just as the live simulated combat exercises has two forces meeting in the real world, the Engagement Response Server can have two forces meeting in the virtual world.

The information that is necessary is not just about your own forces or the terrain they occupy, but also concerns the opposition forces. In a Engagement Response Server the notion of the foe must be considered. All the questions, parameters and capabilities that the Engagement Response Server has for dealing with your own forces must be available with respect to the foe. It is up to the Engagement Response Server manager (or commander in charge) to determine what information can be made available to the opposing participants regarding each other in the Engagement Response Server, but providing thorough virtual reality capabilities is necessary. Thus, for example, one side may have knowledge of the opponent's location and size, but not of its plans and logistics information (which may actually be residing on the same hard disk drive, but not logically available).

Intelligence is of key importance to the Army. Its quality and timeliness can mean the difference between success and failure. Since intelligence is very often incomplete, contradictory, incorrect, and of questionable reliability, a means of validation would be useful. A useful Engagement Response Server would permit not only a means of using intelligence, but also means for analysis and review.

Levels of Informational Needs

One of the requirements for a useful Engagement Response Server is that it operate at different levels to provide the necessary information and capabilities appropriate to the level of the user. Bluntly stated, different users have different needs. The informational needs of a battalion-level commander will be drastically different from the informational needs of a tank commander. Each is valid for its own purpose, and each is necessary for a complete army to function. It is necessary for a Engagement Response Server to provide different varieties of information each tailored to meet the requirements of the appropriate levels of hierarchy found in the Army.

In addition, different commanders at the same level may have different tastes in information. One battalion-level commander may want only the overall picture, while his successor may have a penchant for micromanagement. This variation in needs and wants is certain to exist at all levels. This is a good example of a function that is considered important, but whose implementation should reside with the Army.

Since Engagement Response Servers are software, and some users may want the capability of easily changing the level of information they get, it can be easily provided. Conversely, the capability of locking out certain users from access to certain information should also be provided. This is envisioned mainly as a feature for the Engagement Response Server operator (or commander).

Realism

The Unit Performance Summary catalogues many failures. Some of them occur because of a misunderstanding of the plan, while others occur because events "in the field" made implementation difficult or impossible, even if the plan was well understood. It is clear that a textbook understanding is well and good, but that actually attempting to do something is where much of the learning actually happens. With a Engagement Response Server we need to be able to "transfer the field into the classroom."

How can we make the Engagement Response Server similar to the real world? This question derives from the need to make the Engagement Response Server more than just a high-level intellectual exercise. Realism can be increased by providing high fidelity visuals and sounds and by including the chaos and uncertainties that are typically present on the battlefield. But we may be asking the wrong question. A better question might be: How can we use virtual reality technology to improve our training in a cost effective manner? The answer to this is quite different than providing a veridical virtual environment. It involves creating a system and actually using it, with an eye toward achieving the goal of improved training. Just as we find that much learning in the real world occurs by actually attempting it in the field, we also find that much learning about virtual reality occurs by attempting activities in virtual reality.

Collecting Data for the Engagement Response Server

Identification of Requirements

Here we focus on collection of visual data from remote sites which can be collected into a virtual reality database that can be then compressed and sent in real time to a field command center for immediate playback for mission rehearsal.

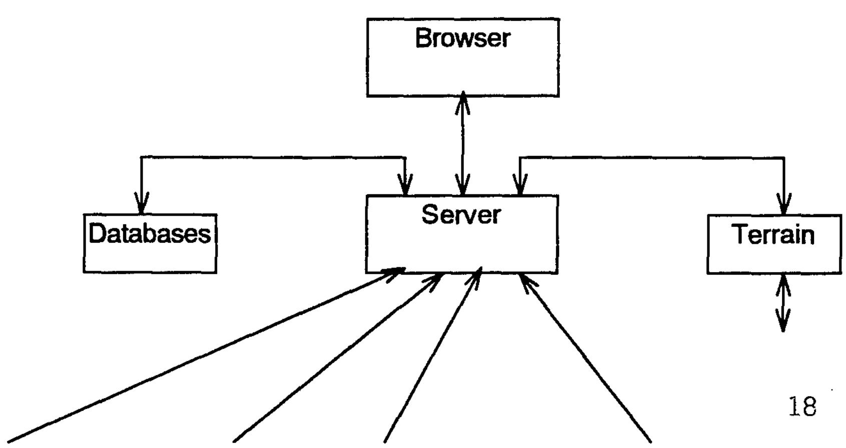
The next area of interest is the integration of visual data with textual and terrain data. In particular, we envision the integration of the Remote Access Controller with GPS and various coordinate systems to enable the integrated presentation of visual, textual, terrain and map data with schedules.

A schematic diagram of the Engagement Response Server is shown in Figure 4, with the video component highlighted.

For visual data, we used both video input, both from video sources such as the Connectix QuickCam camera as well as the newer digital cameras now appearing on the market (Kodak DC-40).

The operation is envisioned as follows:

- The data from these sources is tied to a registration device (see below) so that the data can be "stitched" together to form a seamless video scene of a particular site.
- The data is linked to audio as well as GPS data.
- GPS an other information is used to integrate terrain and map data.
- Hotlinks are embedded into relevant scheduling and textual information as well as other databases from existing legacy systems.



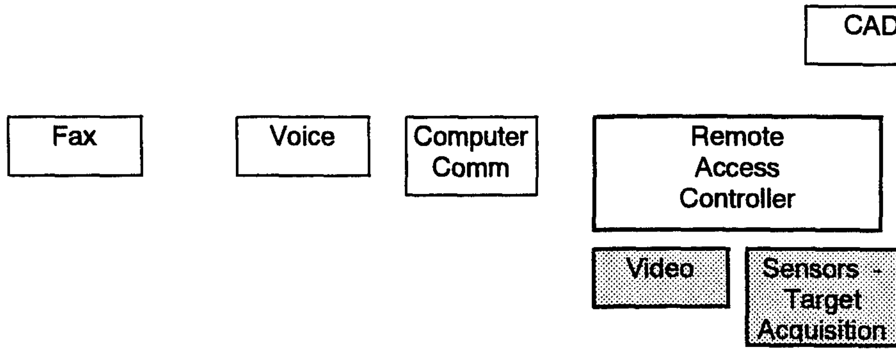


Figure 4: Engagement Response Server Overview Highlighting the Video Component

- Jumps or jackins are embedded to other relevant immersive mission virtual reality scenes. Jumps are locations/objects in the virtual environment which can be "inhabited" by the user. The user can jump from location to location (viewpoint to viewpoint) throughout the virtual world as desired. For example, a commander can inhabit the virtual world from the perspectives of his subordinates, or perhaps take God's eye views from strategic locations. Please note that it is perfectly feasible to locate a jump attached to an object or another user so that when you "jump" to that point you are in effect seeing the environment from the point of view of that object/user. So jumps/jackins need not be attached to static locations.

Registration Device

Currently the means of creating panoramic photographs for immersion require taking numerous shots with a still camera (digital or film) by panning the camera on a tripod, then using a software package such as Adobe PhotoShop to stitch all the individual shots together to create a seamless panoramic scene. A common camera can be used to take taking numerous shots by panning it on a tripod. Accuracy is key here, so a quality tripod is recommended, one that can be level as well as precisely index, or register, the camera from one shot to the other. It is important that the registration be accurate, or the panoramic image will be distorted. A registration device allows the camera to be accurately rotated (panned) a set amount of degrees to allow numerous overlapping shots to be taken.

We propose to join both these technologies together in a registration device. The registration device would be small and portable and easily attached to the Remote Access Controller. Together with the Remote Access Controller and a digital camera it would provide the ability to correctly registering the panning of the camera (or other sensing device) and also stitch together the shots to automatically generate panoramic files necessary to create the mission rehearsal virtual reality scene.

Uses

A sample benefit of this capability is that through the use of off-the-shelf equipment such as digital cameras, a team can be sent to various sites and a database of possible mission sites can be built up. Since these sites are linked with regular CAD and terrain data such as AutoCAD DWF files, existing data from architectural, construction and other files can be added.

Constructing a Mission Rehearsal Database

It is envisioned that the user will be able to create a mission rehearsal database from components of already existing virtual reality mission databases on file. However these virtual environment databases can be difficult to navigate. It is not easy to identify adjacent spatial structures or how to find desired ones. Large libraries may be arranged in an alphabetical file structure and documented by descriptions that do not represent their contents well. These virtual environment mission libraries are often difficult to scan and audition easily because they are not organized according to well identified qualities. It is time-consuming to browse the database directly. The requirement is to identify the components of a mission database, and classify them so that it is easy to rearrange components to provide new training or mission rehearsal scenarios.

Databases

Existing databases can be integrated into the Engagement Response Server to link information to virtual objects or behaviors. The Engagement Response Server can be more than a collection of colorful (or not-so-colorful) graphical objects in a virtual environment. The Engagement Response Server can be an interface to information that is associated with those objects. Such information may be very simple and directly associated, such as the identification number and fuel status information for a particular tank. Or the information may be subjective or abstract, such as textual information describing a certain commander's opinion on some activities. Since the information in the database may be numerical or textural, as well as extensive, it may not be desirable to have it appear visually in the graphical virtual environment. Just because the data is there, it doesn't have to be shown, only to clutter up the field of view. The database information can be accessed only when desired.

The database can be dynamic. That is, the existing information can be changed or new information added. This can allow comparisons of actual performance versus planned. In addition, it can provide a means of linking and understanding intangibles as well as concrete actions.

Avatars and the Engagement Response Server

A key component to immersive visualization for mission rehearsal is the ability to communicate with others in multi-participant virtual worlds (or via the virtual world). This makes it necessary to have representations of people within the virtual world. These virtual people, or avatars, are discussed in this section. An extension of these "people avatars" are avatars for the Engagement Response Server components. Thus we envision avatars for the Internet Appliance controllers. Please note that the use of avatars for objects is non-standard usage. In our discussions we emphasize that avatars may be persons or Engagement Response Server components.

Definition

An avatar is an animated character within a 3D virtual world that represents the user or Engagement Response Server component. You live through your avatar, which you control however you choose, and can customize it as you please. Avatars respond to the inputs of the user, for example, by walking and talking. They may also provide data, e.g. GPS location data.

This distinguishes them from a host of other active or animated objects such as autonomous agents, bots, digital biota, synthetic sentients, and other self-controlled creatures of digital space.

Typically you see your avatar from a "God's eye" view, walking about 10 feet ahead of you. This permits seeing how your avatar interacts with other avatars. This is different from visualizing the world from your own viewpoint (i.e. through the avatar's eyes) as is common in "traditional virtual reality."

Avatars have several dimensions:

- Philosophical Dimension
- Human Dimension
- Technological Dimension
- Representation. All aspects of an avatar's appearance
- Behavior. The dynamically changing state of the avatar
- Interaction. The ability of the avatar to affect its environment and be affected by it
- Identity. The relationship of an avatar to a real-world individual or organization.

Avatars have persistence of identity, so that your avatar retains the customization you have done. Also, they can retain information and qualities from previous activities and interactions. It also has the same representation to others, so they do not see something different each time they encounter you (unless you have explicitly made changes since your last meeting).

The idea behind using avatars to represent persons as well as components of the Engagement Response Server is to facilitate ease of use and to engender a sense of community. Although we have no empirical evidence, we feel that by presenting a "human" face and embedding control within a "human" context, the user or soldier is able to personalize the control of the Engagement Response Server to the extent it becomes a trusted tool. Subconsciously it becomes like a pet dog which provides help and support as needed. The other soldiers (as avatars) help also to develop a sense of the larger community which is important in the coordinated defense service missions (joint Army, Navy, Air Force operations).

In addition to the above mentioned abilities of facial expressions, limb movement, and locomotion, avatars have other capabilities:

- *Tour bus capability.* Other avatars can be linked with your avatar and automatically go with you and see what you see, for example, jumping from hyperlink to hyperlink
- *Collision detection.* Can be on, or off to pass through objects
- *Physics.* Real world attributes such as gravity can be manipulated
- *Intelligent agent technology.* Adds behaviors to avatars, such as walking, other limb movement, ability to respond to activities, such as seeking a moving target, or search for information
- *Physical Interactions* representing something you can do in real world (handoffs, combine effort to push, pull, carry objects)
- *Virtual Interactions* that cannot be done in real world (e.g. clicking on an object to access a database)
- *Searchability.* Allow information attached to avatars to be searchable, vis standard internet search engines
- *Security.* Avatars can be assigned public keys or certificates to enable secure communications, online purchases, and transparent micro-transactions while in a virtual world. This can also include authentication and verification information.

The Sense of Community

The most popular current use of avatars is in online chat rooms for socializing and recreation. Examples are WorldsAway, Palace and America Online. People from around the world use their PCs and modems to access specific virtual worlds to meet and interact. These virtual worlds are typically rich 3D environments that lean toward fantasy or science fiction motifs and usually have a theme, such as a hobby or special interest, or perhaps some kind of architectural significance. The worlds are typically limited to a small number of avatars at once, in the range of a dozen or so. But since there are many of these worlds, the users jump from world to world to explore, to have fun, or to find a new stimulus.

The software that enables the worlds and avatars typically permits great latitude in customizing the avatars. The user can easily change their avatar's characteristics, such as shape and color. Nowadays typical avatars are cartoon representations of people, but any kind of object is possible, such as geometric shapes or animals (butterflies, rabbits, etc). Since the virtual worlds are often run as a profit-generating business, the company that runs them may have certain constraints, for example, avatars must be of certain sizes, or they must be selected from a stable of avatars (which can be then customized by color, clothing, type of head, etc). The avatars are dynamic and respond to their users' controlling inputs, and are capable walking, smiling, frowning, waving, jumping and gesturing.

These worlds are often offshoots of the online chat rooms and communication is key. The basic form of communication is still the text-based online chat, with the words attributed to an avatar appearing in a cartoon speech bubble nearby. Today's simple chat room type of virtual worlds have the look of animated Sunday newspaper cartoons.

Advancements are rapidly being made away from simple cartoon-like chat rooms. New developments permit the use of voice and graphics. By using a sound card, microphone, speakers and appropriate software, one can speak and hear instead of having to look at text-filled balloons. With this audio capability the worlds become more realistic and the display of text on the screen, which occludes the view of the world, is eliminated. In addition to audio for voice communication, the worlds can have 3D audio capabilities in which sound appears to emanate from a virtual object.

Much work is also being done in the way of improving the visual aspects of both the worlds and avatars. Avatars can now be texture mapped, allowing your avatar to have your own photographic likeness on it. The worlds are also becoming rich in texture mapped graphics, rather than simple shaded objects. Panoramic surround video can be used as a backdrop, with the avatars or other 3D objects placed in the center.

Although simple chat room communities exist, mainly for entertainment, they can be extended using off-the-shelf software and hardware into the Engagement Response Server. These extensions can still be viewed as communities, in which the communities are the people who are training or rehearsing missions. The term community does not necessarily mean only those who are on a specific training task. For example, if a platoon was rehearsing a mission, the platoon, the adversary, non-combatants, as well as the invisible observers (remotely located commanders) can be said to comprise the community. Virtual reality is being extended to include multiparticipant environments, which implies that there is a sense of community, whether it is to a lesser or greater degree.

One common theme running through our reports has been ease of use. We feel that one of the key advantages of a virtual reality based system is the relation to common physical functions making its use easy. One reason to use the internet paradigm was that large numbers of persons are being exposed on a daily basis to the internet and more will be, especially with the move to the common PC/TV internet devices. It is becoming a new kind of community. Ease of use is important since many of the users will be soldiers and others who have limited time and in emergencies have no time yet must understand and use the Engagement Response Server to rehearse important missions. Ease of use combined with readily available access helps to engender a sense of community.

VRML - Virtual Reality Markup Language

The goal of VRML is to create a seamless distributed multiuser virtual environment that connects all the 3D worlds on the internet. VRML, along with its extensions, provides a complete distributed multiuser communication infrastructure for complex virtual worlds. VRML is an open specification that has been embraced by Netscape, Microsoft and Silicon Graphics, essentially making them available to every browser. Several high-quality 3D VRML scene and object creation tools are now available, and manufacturers are now competing to provide the fastest, most robust, and most standard-compliant VRML viewers. The trend to open standards, as opposed to proprietary technologies, is continuing with VRML. This will allow content developers to create multiuser worlds that run on any server.

VRML allows much more than walkthrough of simple 3D virtual worlds. VRML integrates multimedia - still video, motion video, text, voice, 2D/3D sound, also databases. Open standards allow for addition of common features, such as shared whiteboards, post-it notes and audio chat.

VRML uses 3D geometrical objects as building blocks to construct virtual worlds. Texture maps or background environment maps can be added to enhance the complexity and realism of the virtual environment. VRML is the new standard for 3D over the internet.

Piaget once said "the ultimate path to learning is life itself." The rich multimedia VRML environments are growing in their ability to simulate the richness and chaos of life itself. We are working to apply this to the goal of this virtual reality database, which is to use off-the-shelf software and hardware for immersive mission rehearsal. By integrating VRML with existing technologies we can expand communication and learning, which is the core of mission rehearsal.

Webcasting / Mediacasting

Webcasting is a shared experience for the Engagement Response Server, extends the broadcasting paradigm and blurs the boundaries between audience and participants. Unlike broadcasting, there is the opportunity for realtime participation and feedback in entirely unprecedented ways. An example is broadcasting live video over the internet, with the audience being able to respond via email.

The Audio Aspect of the Engagement Response Server

The use of sound has not been explored very thoroughly so far in this system. However that is not to say it is unimportant. In fact the use of sound to provide another dimension to mission rehearsal is important. Use of sound extends far beyond just voice and voice annotations. The basic auditory interface techniques of auditory icons, earcons, and sonification need to be explored as to how they can meet the needs of a mission rehearsal system. A part of the Engagement Response Server must cover the hybrid auditory interface techniques and methodology for creating hybrid auditory interface techniques for use in a complex mission rehearsal system. In particular, its use in a distributed networked collaborative virtual environment.

Luckily audio research is moving in the direction of the internet. In particular, the use of sound in the collaborative, networked mission rehearsal environments. For example sound can be used to navigate the virtual mission database. Sound can be visualized as lines linking each mission component to its child and parent nodes, spread across a plane in front of the user. A user can browse by following the visual links between groups, traveling from an auditioned component to other components. If the user moves to a new group of sounds or presses the selection button, the closest objects are sounded and change color.

Many of the above topics were discussed at the latest International Conference on Auditory Displays 1996 which was held November 4-6, 1996 in Palo Alto, California. The key topics in this conference relevant to this project were the use of sound in the context of virtual environments and the internet. Conference topics were :

- Auditory data representation (exploration of data via sonification)
- Sound in immersive environments (virtual reality)
- Auditory displays on the World Wide Web.

An Interactive Collaborative Auditory Environment on the Internet

We are interested in creating an infrastructure for providing an interactive collaborative auditory environment on the Internet So we propose to add a parallel data stream to the visual data stream for the engagement response server The streams share some common control and synchronization mechanisms and information sources. The output of the streams is what the user will hear and see while using the system

The particular areas we are interested in are exploring the ways to provide :

- A data sonification toolkit or sonically enhanced Engagement Response Server
- An auditory display schema suitable for mission rehearsal
- A 3D-auditory environment for hierarchical navigation in non-visual interaction
- Ways to provide an interactive collaborative auditory environment on the internet.
- Embedding of voice/audio data.

There are several ways to use voice and audio data. One is to embed audio files into mission rehearsal scenes and to broadcast audio over the internet (live or pre-recorded). With this, sound effects or audible commands can be linked to virtual objects and events. Also, voice communication among users (avatars) makes live, real-time communication feasible. This audio information can be spatialized to provide 3D directional sound, which would be important in presenting the user a more realistic virtual environment. The voice communication among the users can be stored for later playback and analysis or archiving.

Whatever is designed has to be extensible because audio and it's uses on the internet is an active research area. However even at this stage there are certain trends that can be used with appropriate modifications. We would like to paraphrase and highlight those areas of active research which we feel can be added to the engagement response server to provide an auditory environment.

Main areas which seem promising to explore. These are :

- Automatically analyzing the information contained within large numbers of documents. For example, we may use a statistical measure of key words computed from the text to represent the semantic content of each document. SPIRE (Spatial Paradigm for Information Retrieval and Exploration) is a system for analyzing the information contained within large numbers of documents, using audio as a means of presenting the semantic content. SPIRE was developed by Pacific Northwest National Laboratory for the intelligence community. SPIRE research was supported by the Joint National Intelligence Development Staff under a Related Services Agreement with DOE under Contract DE-AC06-76RLO 1830, Interagency Agreement Contract No. 21607A.

- The expansion of work being done by Leonard N. Fone at the MIT Media Lab on Wearable Augmented Sensory Systems. This work is directly relevant to mission rehearsal because it uses data sonification to compensate for normal limitations in the human visual system to identify visually indistinguishable objects. The following is a quote from Foner's works.

  Almost any three wavelengths of light, if their amplitudes are adjusted correctly, can be made to look like some particular color so that the human eye is easily fooled. Basically the work involves imaging in 128 wavelengths in the visible spectrum (and, optionally, another 128 wavelengths in the near-infrared), then using sonification to make the resulting histogram of wavelength amplitudes accessible to its user. So sonification is used to keep the user's visual field uncluttered.

  The system is wearable. It sits on the side of the user's head, and images a patch of his or her environment about two degrees wide (the same width as the fovea), aimed in the direction that the user's head is pointed. (Full eye-tracking is too intrusive and too cumbersome to justify.) This makes it possible to use the system most of the time, and to *learn* the mapping between materials commonly encountered and their sounds. This in turn may allow the user to identify what things are made of, or whether objects have undergone a change. (For example, imagine looking at a car and saying, "Well, it *looks* like metal, but *sounds* like painted plastic," or looking at one's lawn and saying, "Hey, the grass sounds funny today—is it sick?"). Future extensions of the system to a wider electromagnetic bandwidth (e.g., far-infrared and near-ultraviolet) promise to substantially improve its utility in that regard, as do designs that incorporate unusual, nonhuman senses, such as a polarimeter, an RF-field sensor, or a magnetometer." (Also perhaps the use of infrared sensors and sonification to help distinguish between similar objects).

Sound Sources

We have started experimenting with experimenting with sound sources. Currently we are experimenting with local sounds which need to be downloaded and also the use of streamed audio (RealAudio, Xing or other network broadcasting tool).

Problems

The major problems with Web based audio are lag and scheduling problems due to packet-based data delivery and sonic quality available via the very limited bandwidth. Limited download time: we vowed to constrain our download time within a 100K limit if possible. Since downloading samples means establishing a new connection for each sound file (which often costs more time than the actual downloading), we had to limit the number of samples as well as their file size.

Limited real-time bandwidth: responsiveness is very often the key to an interactive installation's success. To achieve this, the nature of the data that flows back and forth should be very low bandwidth (< 1 Kbytes/second).

Unreliable UI (user interface) events timing means that literal mapping is prohibited. real time responses to changes in the environment.

Enhancements to Browser

It must be understood that the functions of the server must have corresponding functions in the browser. For the sake of simplicity in this report we assume that when we discuss the server, we are also covering the browser.

In many cases the functions of the server are set up during the installation, and these may be placed in the browser instead.

Enhancement: Sample Translator for CAD Data

Standards - Drawing Classification System

Virtual Reality Database Standardized Numbering System

By organizing virtual reality databases in a systematic manner we can simplify the tasks of finding information and quickly understanding the contents of files. The key elements of a standardized numbering system for a virtual reality database concerning the drawings portion are:

- Virtual Reality Database Name
- Virtual Reality Database Description

- Virtual Reality Database Identification
- Virtual Reality Database Manager
- The application software interfaces (legacy system interfaces) to be used throughout the entire virtual reality database.

Virtual Reality Database Directories

All virtual reality databases will be accessible via customized interfaces. The interfaces will give user access to all virtual reality databases owned by the same group. The level of access to all virtual reality databases is controlled by field commanders and above.

File Organization

Introduction

This section provides an overview of strategies for the organization of CAD drawing files for a given virtual reality database, user group, application software product. More detailed information is given in Appendix D.

Reference Files

Reference links are very important in any classification system. It is envisioned that only the most essential data is kept by the Engagement Response Server. In keeping with the internet web philosophy, dynamic links will be used for example to dynamically link a drawing to one or more other drawings. The drawings themselves may be kept by the responsible organization. Another value in doing this is that information can be distributed to many drawings in a virtual reality database and as the base information changes each of the linked drawings is automatically updated.

Filename Format

A specific naming convention should be adopted and then followed by all user groups. Failing to do so may have adverse effects when referencing a drawing. Such a convention may be hierarchical, with information being represented from a general group down to a specific group. For example, it may include tags ranging from Field Command to Architectural to Site Layout to Revision Number.

Enhancement: Terrain and Maps

The classification system above can also be extended to include identification of appropriate terrain and area maps. It can be used to manage the placement of overlays on the maps and provide a global map structure to point to other information for mission rehearsal. One area which of course is necessary is to cater to different coordinate systems. See "Technical Reference Datums, Projections, Grids and Common Coordinate Systems, Transformation of", Department of Defense, MIL-HDBK_600008, May 1991 DMASC-SGS, 8613 Lee Highway, Fairfax, Virginia 22031-2137.

The Internet Appliance

The work on the Internet Appliance is an outgrowth of the Remote Access Controller. As such, we begin with a discussion of the Remote Access Controller.

Remote Access Controller

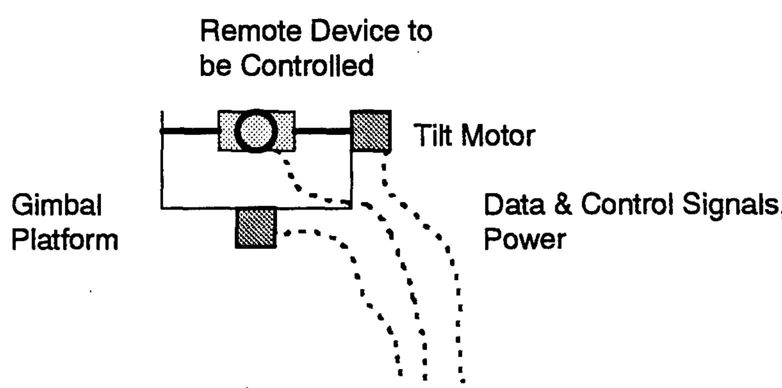
The Remote Access Controller consists of a communication component and a motion platform. The communication component interfaces with the internet using its own protocol to provide data and obtain control information. The controller protocol consists of simple ASCII commands.

The Remote Access Controller motion platform consists of a tilt/pan mechanism actuated by stepper motors, Figure 5. One motor controls tilt (up and down, pitch) and the other controls pan (left and right, yaw). The motors are driven by a controller that provides both power as well as control signals.

There are a couple of ways of setting up the remote site depending upon the needs of the user. The description above uses standard PC hardware and software. It is possible that the functionality of the PC and its software could be placed on a tiny PC such as is used in embedded control applications. With further integration, a custom controller can be built that includes all the PC and modem functionality.

Additional Features

This Remote Access Controller is a basic system that can be extended as additional needs and capabilities arise. For example, the controller may be pre-programmed to follow a specified path over the course of the day, or perhaps a more intelligent program can be used to "lock on" and follow a target.



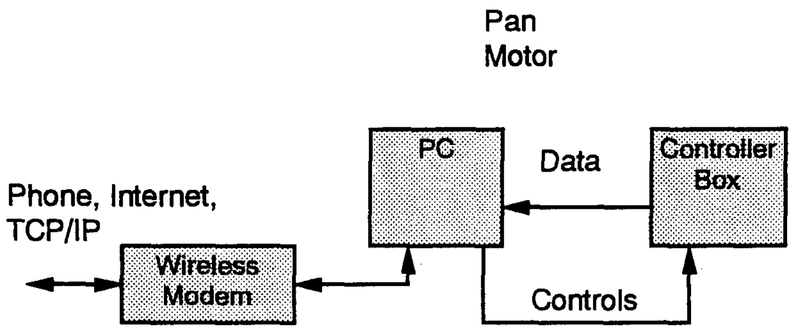


Figure 5: Schematic of Remote Site

Archiving for Mission Rehearsal

By tying the Remote Access Controller with GPS and map capabilities, the data retrieved by the Remote Access Controller can be stored in a database and later be used for mission rehearsal or post mortem analysis.

The Remote Access Controller could even be set to operate automatically, thus enabling the capture of a series of data shots as the situation unfolds, to create a database that can be used for later playback. A concrete example of this archival ability is later discussed under as an application of the remote access video controller.

Remote Access Virtual Presence

Nowadays it is possible to send live video over the internet. It is considered somewhat of a leading edge technology today but the cost of the equipment is dropping and the technology is rapidly losing its avant garde edge. Video cameras that plug into the parallel port of PCs are available at computer stores. These devices are targeted at users who want to capture video snapshots, record video streams, or connect to the internet for broadcasting live video. Our concept builds on this off-the-shelf technology.

The video component will provide the ability for any one to be able to "plug in," or in the words of the science fiction author William Gibson in *Neuromancer*, to "jack in" to any mobile platform and experience the situation that that platform is in.

A picture is worth a thousand words, so the ability of a field commander to actually visualize any situation involving any of his assets/mobile platforms such as personnel and equipment is invaluable.

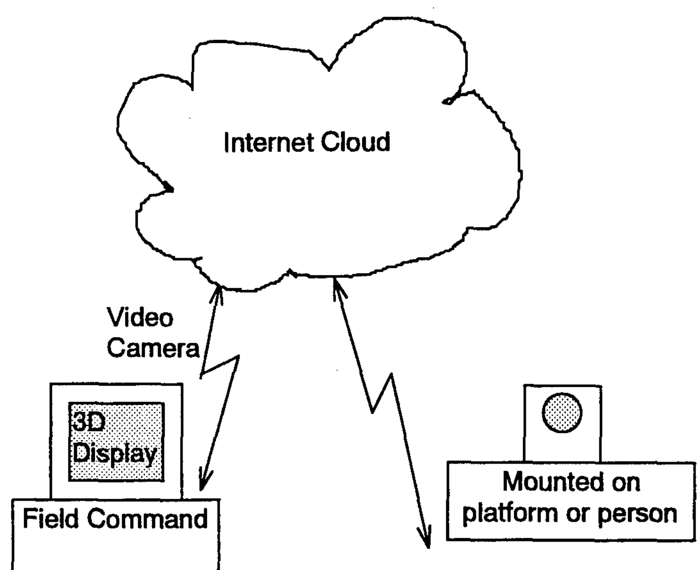
We provide a cheap low cost remote access video camera which can be controlled remotely by anyone with access to the engagement server to view and control the camera from a remote location. The key element is to be able to control the camera remotely using a virtual reality device such as a Virtual I/O tracker. Using a head mounted or "CAVE" like display device, the remotely based operator is able to move the camera left and right, up and down.

As shown in Figure 6, the implementation has two components, local and remote. At the local site is the software and the input device for control (such as mouse, HMD/tracker, joystick, etc). At the remote site is the motorized camera with its capability to tilt and pan. In operation, the user would see on his screen or on his head mounted display the picture taken by the remote camera. If the user has a mouse for control, he can use mouse actions to cause the camera to swivel left and right, up and down. If the user uses a head mounted display and tracker the camera would mimic the motions of his head. So if the user tilts his head upward, the camera would be driven upward. The video is sent to the head mounted display, so he sees what the camera captures. This accomplishes virtual presence over the internet. A camera with motorized zoom would respond to zoom in and zoom out commands. Since this is a internet application, the views from the camera can be made available to everyone on the internet or kept restricted to those with a need to know.

Local Site

The software and hardware at the user's local site is based upon commercially available standards, such as a PC compatible computer linked to the internet with a modem and phone line. The software running on the PC would be a generic operating system such as Windows95, along with along with a generic internet software or web browser, such as Netscape Navigator or Microsoft Explorer.

Our design philosophy of designing our software/hardware to piggyback onto existing software/hardware provides additional functions at no cost. As new functions become available from manufacturers, the engagement server automatically makes use of them. A concrete example of this is that since we started exploring the remote access video, Connectix (one of the video cameras we use) has come out with beta software for whiteboarding and video conferencing. Thus the remote access video option now can make use of these functions.



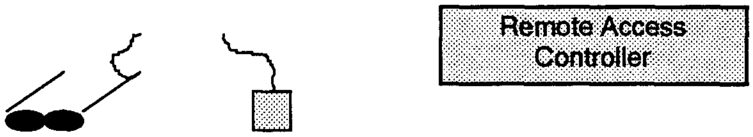


Figure 6: Schematic of Remote Access Video System

Our software for interacting with the remote camera would be installed and running, and interactions would be with a mouse or with a head mounted display with tracker, such as the Virtual I/O i-glasses. With this suite of hardware and software the user could then control the camera at the remote site.

Remote Site

The components at the remote site allow the inputting and transmission of live video. This is accomplished with a combination of hardware and software, in particular, a PC with a video camera, motorized camera platform, a modem and attendant software. Once again, commercially available parts are readily available for this. The camera could be one of the commercially available devices (such as Connectix QuickCam) designed to easily install on a generic PC. Video cameras of this kind typically require a PC running Microsoft Windows 3.1. In addition, a modem and modem software is needed to communicate to the internet and modem software is also required to communicate to the camera tilt and pan motors.

The remote PC has three functions:

- Connectivity to the internet,
- Acquiring and transmitting video signals, and
- Controlling the Remote Access Controller motion platform.

Additional Features

This remote access video system is a basic system that can be extended as additional needs and capabilities arise. For example, the camera may be pre-programmed to follow a specified path over the course of the day, or perhaps a more intelligent program can be used to "lock on" and follow a target. In addition to merely displaying the video stream, it can be archived in many different ways, for example, onto hard disk, tape or burned onto CD ROM. Not all visuals need to be veridical, so false-color images can be also viewed (e.g. temperature gradients) or even night vision capabilities provided.

Wireless capabilities would also be very important to untether the remote site from phone or communications lines. This would be mandatory for mobile applications.

Since the issue of secure transmissions will undoubtedly be important, means of encryption of the video and audio signals are envisioned. These features enhance the basic remote access video system to provide an integrated hardware and software solution.

The ability to access the video/audio signals causes no contention among multiple users, since viewing and listening involves receiving copies of information. So a benefit is multiple field units can receive the video signal helping communication problems. Note however that the ability to control the camera is limited to a single user at a time (or else a well thought out sharing scheme designed). It is important that user have a coherent session of camera usage or else sensory dislocation will set in, resulting in frustration and annoyance.

Applications

This system can be used in different types of remote locations and applications. The camera can be permanently mounted, transportable for temporary setup, or mobile. A permanent installation is one that would be expected to remain in place for years, for example, a camera available to the community to remotely view the local conditions of highway congestion. Size and weight would not be much of a concern for a permanent installation. A permanent installation would probably be hard wired into place.

A transportable system could be used where frequent moves are needed, but size and weight are not essential. Wirelessness may or not be a necessary requirement. It may even be vehicle mounted, for example, mounted on a tank. A civilian application could be for a construction company to review progress and identify problems at a construction site, and can be moved hourly, daily or weekly as the project progresses.

A mobile system would be small enough to permit a person to use it while on the move while also carrying other equipment. A mobile system seems best suited to the purpose of the infantry and would need to be small, robust and lightweight.

Example of Archiving for Mission Rehearsal

As the mobile platform, e.g. a soldier with a remotely controlled camera, progresses through the terrain, the camera can view and store the video images together with position information. The position information can be taken from a GPS unit and correlated with map data. These video images with the position and map data are stored and processed via QuickTime VR from Apple or similar technology to provide a navigable video database of any complex situations. These databases can be used later for mission rehearsal or post mortem analyses.

So the remote access video enables the capture of a series of panoramic shots as the soldier progresses, to create a database of terrain scenes that can be used for later playback via immersion. A remote user could then do a virtual reality walkthru of the same terrain as the soldier in the field.

Implementation

We took the first steps toward implementing the ideas discussed above and constructed a working camera and Remote Access Controller that displays video images on a PC. This system consists of a PC that runs the video software and a second PC that controls the motion platform of the Remote Access Controller (Figure 7). This separation was done to simplify the tasks and allow the basics to be tackled first using available equipment in our

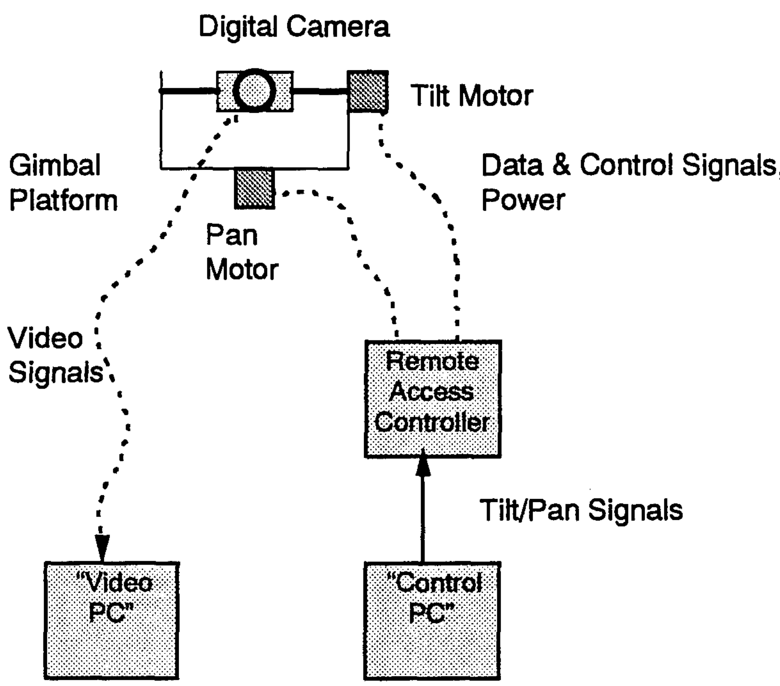


Figure 7: Implementation of Remote Access Video System

lab. It allowed complete de-coupling of control functions from video play functions. We constructed a simple motion platform in our lab which included the controller hardware and software as well as a gimbal mechanism with motors. This device was sent to the Army for evaluation.

Usage

This implementation uses the video software completely separate from the control software. On the "Video PC", the video program is executed under Microsoft

Windows/Windows 95 and the video window displays what the camera happens to be pointed at.

Remote Access Controller Protocol

The "Control PC" runs DOS, and a simple modem program, Procomm, is executed. With the appropriate port setting (Com 1) and baud rate (2400, N, 8, 1) we are now ready to enter keyboard commands to control the motors.

Commands consist of a number and a letter. The number is the amount of "ticks" the stepper motor will index, each tick being 3.75 degrees (96 steps per revolution). The letters that control motion are l (left), r (right), u (up), d (down). The protocol is to enter a number, a space, and a letter. A "1" is returned to the screen to indicate a correct response by the stepper, and a "0" is returned to indicate a fault.

In addition to the commands to move the steppers, additional commands are available. To change the time in between steps "w" is used. To set a new zero point "s" is used, to move to the zero point "z" is used. To query for ready "q" is used. A zero point can be selected, for example, to view a certain object. Once set, that point can immediately be recalled by using the "z" command.

Microcontroller Hardware

The Remote Access Controller is based on a commercially available programmable microcontroller with the addition of driver circuitry for powering the motors. For the programmable microcontroller we used the BASIC Stamp from Parallax Inc, a low cost stamp-sized computer that runs BASIC. It is comprised of two main components, an 18-pin BASIC interpreter and a 256-byte EEPROM that holds a tokenized version of our BASIC program, which is read and executed by the interpreter. The remainder of the Stamp is taken up by a 4 MHz resonator, 5-volt regulator and 9-volt battery clip. Also included is a small prototyping area which provides connection points for the 8 I/O lines, 5-volt supply, unregulated supply and ground.

The BASIC stamp is programmed in Parallax's simple BASIC language. Programs are written on a PC in the included editor, and a programming cable permits downloads via the PC's parallel port. Since each instruction takes about two or three bytes, and there 256 bytes of EEPROM space, maximum program size is limited to about 80-100 instructions. Speed is approximately 2000 instructions per second.

Remote Access Video Control

The issue of who is in control becomes important. This is much more than doling out passwords to control a remotely controlled camera. The notion of passing control to another specific person becomes important. For example, if a soldier is using his browser to remotely control the camera to view something, and he needs to pass on control to his commander, there must be a means of communicating several things. First is the need to tell the commander to get ready to assume control. Then comes the need to pass off control. Next, it may be necessary for the recipient to acknowledge that control has been received. This is to avoid a problem where the first user passes control to a second, but the second has a problem or is unable to take control, leaving the system unused and the potentially valuable information lost.

Since the remote access video can be accessed simultaneously by multiple field groups and even from one mobile platform to another, control coordination is important. An instance of coordinating control may involve scheduling among a group of users.

The Network

Another fact to be aware of is the requirements imposed by a particular kind of network. In our case, we are using the internet. The goal is for a user to be able to remotely sense live video with his internet software. This means that the remote sensing equipment must be sending signals up to the internet, with the implication being that the appropriate network interfaces are in place. One solution is to have the remote sensing device (the video camera and motion platform) connected to a PC which runs the appropriate operating system and video applications programs as well as the appropriate motor control program as well as the network program. In many cases this may be a prohibitively bulky and delicate solution; a small robust and mobile solution may be required. In fact, the best solution may be to build an integrated system housing all the computer and communications equipment in the motion platform.

The realities of the physical world become very apparent when constructing a device that is to be used remotely. The device may or may not be attended, but in either case, it must be rugged enough to be used in its intended environment.

The device should be designed such that if power is removed and reapplied, that the camera will reset itself by pointing to a known zero point. This may be "straight ahead" or may be at a certain object. Because of memory limitation we omitted this, but it makes starting a new session much easier and understandable.

Evaluation of the Remote Access Controller

We shipped the Remote Access Controller to the Army Research Institute's Simulator Systems Research Unit for a week's evaluation. The purpose was to show the preliminary implementation of some of our ideas to better convey the direction we are headed. It must be kept in mind that this particular device is an engineering prototype and the intent is to evolve it into a more useful and robust device. The Remote Access Controller was then returned to us to continue development.

A user's manual for the Remote Access Controller was written, and for a starting point the device and software is known as Version 0.1. This user's manual provides the basis for succeeding versions and is intended to be continually updated as required. The latest user's manual, Version 0.2 is included in Appendix B.

Further Enhancements

We plan to move forward toward full remote control as well as improve the mechanical and system control aspects. For the Local PC this involves the host software. For the Remote PC it involves software to receive commands over internet and then send them to controller box. In addition, we see that zero point at startup should be added so that the camera will always face the same point upon startup. Also, the capability of bookmarking coordinates, that is, to save coordinates of interest would be a welcome addition.

Another addition that would be useful would be a video camera emulator. This would provide the capability of a Virtual remote camera that uses VRML or virtual reality to simulate a real environment.

The other component defined further is the terrain and map components. To date we have not seen much development on the net providing the ability to present terrain data on the internet. We will look at some of the issues surrounding the terrain component. Related is also the issue of control and presentation of maps under the engagement server.

We will also like to examine various compression formats such as the new DWF compression format used by Autodesk. The DWF format is used to present large amounts of CAD data on the internet. The integration of the terrain/map component with the WHIP plug-in from Autodesk and VRML readers will provide a ready made environment on the internet for "immersive visualization of missions."

Off-the-Shelf Video For Mission Rehearsal

We looked at the commercially available items that would provide an immersive environment with video for mission rehearsal.

Surround Video Technologies

The term "surround video" is a generic term for those computerized technologies that give the user the ability to "look around inside a photograph." By using a camera to record a panoramic image of the real world and placing that image on a computer screen, the user can then control the portion of the scene that is displayed. In essence, the user is at the center of the panoramic view, and can see a slice of it on his screen. The four leading products are Apple QuickTime VR, Microsoft SurroundVideo, RealSpace RealVR, and Warp VirtualTV.

QuickTime, SurroundVideo, RealVR and VirtualTV are similar in that the images originate from a panoramic camera or from a series of photos that are later stitched together to form a panoramic photo. Panoramic images allow the user to look left, right, back and forth, and only up and down to the extent that the image was captured by the camera. Warp's VirtualTV can also take a 180 degree wide angle photo (from a fisheye lens) as the source image. By dewarping the image the user can view a complete hemisphere, which includes up and down.

Those products that do not use the fisheye lens approach must capture a panoramic image, and there are two common approaches. One is to use a special and expensive panoramic camera whose lens revolves around a cylinder of film. The other is to use a common camera and to pivot it about a specific point, taking numerous shots that are later stitched together. Accuracy is key here, so a quality tripod is recommended, one that can be level as well as precisely index, or register, the camera from one shot to the other. It is important that the registration be accurate, or the panoramic image will be distorted.

The viewpoint of the user is restricted to the viewpoint that has been defined and processed by the creator of the walkthrough. This viewpoint is where the camera was located when taking the panoramic shot, and this point is called a "node." Since we are dealing with a 360 degree panoramic photographic image, rather than a virtual world comprised of 3D objects, it is not possible for the user to traverse through the photograph to freely and actively explore the environment.

In order to provide the user with the capability of traversing through a photographed scenario, a group of panoramic images can be made, each from a different node. By jumping to different nodes, the user can effectively traverse through a photographic scenario. He is still restricted to viewing the scenario from the nodes defined by the creator, but if there are enough nodes, he can have an increased capability to actively explore the environment.

For example, if a single panoramic shot is taken from the center of a city street intersection, the user can see the buildings along the streets and see each street receding into the distance. But he cannot move down the streets, or move to get a better view of a building. But if a series of panoramic shots were taken, say at nodes 20 feet apart, down the center of a street, all around a city block, the user could traverse around the block by jumping from node to node. In this way he could actively explore the environment of the city block, although he is constrained to the node points spaced 20 feet apart.

We propose to use both digital as well as video cameras to provide surround video capability. The extension we propose is to be able to use a digital camera to capture a series of digital photographs and automatically edit these photographs together with relevant spatial coordinate data to form an "on the scene" virtual reality mission rehearsal database. This database can then either be transmitted in real or later to the field command center.

The extension we are interested in is to be able to build up a series of virtual reality sites of possible future conflict areas which can then be embedded with hooks so that in future if conflict should arise the infrastructure of the mission rehearsal database is already there in a mission rehearsal library. At the same time, capture of data must be simple, cheap and easy. The digital camera (from Kodak, Casio and others) is a key extension of the video camera for use of the surround video technology.

This is because the digital camera provides an easy, cheap and efficient way to capture a series of photographs at a mission site. The very nature of the digital camera means it is inconspicuous, the snapping of pictures by tourists is very common. This enables a panoramic series of pictures to be captured even a key strategic locations. By then tying these photographs with "stitching" software, GPS location and registration data you then have an instant mission rehearsal scene.

Of course, there is more that we can add to this mission rehearsal scene than just the ability to look around in a photograph. An off-the-shelf capability is interaction. Hot spots can be defined on the images, with actions occurring if the hot spot is selected. For example, if we have a panoramic image of a group of buildings, each building may be defined as a separate hot spot. By clicking on a certain building (i.e. clicking on a hot spot) with the cursor, the user can then find out information specific to that building. The information may come in the form databases, schedules, documents, sounds, text overlays, animations, and so on. In addition to application programs, URLs to data from legacy systems from different battle groups can be defined.

In addition to being able to look around inside a still photo, full motion video is also possible. Currently the technology requires the use of panoramic cameras or multiple shots from a pivoting camera, a full motion video sequence is generated by laboriously taking the panoramic shots at many locations along the trajectory. This involves setting up the tripod and camera, taking the panoramic shot or series of shots, relocating to a new spot and repeating the procedure. This involves a lot of work, but the result can indeed be a full motion video.

Warp's VirtualTV, or VTV, on the other hand, easily permits full motion video. By using a video camera with a fisheye lens, and moving it along the desired trajectory, a video file or tape of the warped images is captured. Processing the series of files to dewarp them then provides an "immersive movie." This is a much easier way of creating full motion video.

These products can also be used to add computer generated scenes to our mission rehearsal scenes. Instead of a camera taking a shot of the real world, screen shots of a virtual world created from CAD and other data can be used. This way users can be provided with a virtual reality walkthrough superimposed on the real scene. But since the view is from a specific point in space, the user would be limited in his ability to view the virtual world. He would be constrained to view it from only those points in space that have been defined and processed by the creator of the walkthrough.

Regardless of the source of the images, they can be viewed over the internet. Browsers can easily be upgraded with plug-ins to view the panoramic images. Head mounted displays can be used to provide an intuitive immersion and control.

We are leaning toward using the RealSpace RealVR as an implementation tool. This product is available on the PC platform at a reasonable price. Importantly, we are working with RealSpace Inc to become developers and resellers.

The extensions we will add to the above video software are:

- Integration of spatial data, e.g. the ability to embed GPS location and other data as part of the virtual reality scene. Not only can a scene from a specific known node be identified, but GPS information can also be added a scene by way of hotspots or visual tags. For example, if a surround video scene is taken, its GPS data can be associated with it, so if someone has a specific coordinate, the panoramic view can be called up and viewed. Also, if tags are embedded into the panorama, the user can click on an object and get specific GPS information about it, or perhaps jump to that node to view things from that viewpoint.

- Embedding of related CAD, textual and other data from various databases. 3D objects can be embedded into the surround video worlds. For example, a panoramic scene of a landscape can be taken and a 3D CAD model of a building can be placed in it to give the user an idea of the result of new construction. Also, since the panoramic scene is in the distance, effectively at infinity, small motions away from the centerpoint (or node point) are possible without detrimental effect. So, the user can then make a circle about a 3D object placed near the node, as long as the circular path and the 3D object's distance from the node are small. This allows a user to navigate around 3D objects while using surround video as a background. Text and other data can similarly be placed in a surround video world.

- Multiple users (avatars). Avatars representing different users can be placed in the surround video worlds. Each avatar is controlled by its user. Since the technology (as noted above) enables 3D objects to be embedded within the panoramic views, avatars can navigate around and about each other, as well as around 3D objects, within this space.

- Overlays of various data. There is a wide variety of data that can be desired by users, and we need to make provision for this. Such information may include 2D data such as text and graphs. It may be 3D data such as 3D objects, with or without texture maps. Or perhaps live streaming data, such as video or stock quotes. This large amount of data can become confusing very rapidly, so means of presenting the appropriate information to the user, or allowing him to arrange its presentation, is important.

Description of the Internet Appliance

What is the Internet Appliance? While working with the Remote Access Controller to control digital and video cameras to create mission rehearsal scenes we came across the requirement to tie these scenes to coordinate data such as from a GPS.

Rather than add the data stream from a GPS and other devices such as registration devices needed for virtual reality databases we decided that a new device whose main function is to tie together data streams from various devices and interface with the internet was needed. We term this the Internet Appliance.

The internet appliance is a PC based standalone box to which can be down loaded programs to do specific tasks. It does not require a human interpreter/scene arranger. It's main function is to automatically integrate the diverse data sources as required by the server. Each appliance has a unique identifier. The server can download the programs as required by the user to the appliance to collect, integrate and transfer the data to the server.

For example the Internet Appliance can bring together video (from the Remote Access Controller) with GPS and time synchronization for immersive visualization over the internet. This extends the visualization of terrain by tying it to the location's coordinates as well as time. Thus a user on the internet can see images of a specific location at a specific time. These scenes can be real-time or archived for later use. Real time use depends greatly on the transmission bandwidth available. Experimentation is necessary using the latest video, sound compression techniques to see if an acceptable frame rate is obtained. Other techniques such as storing and later forwarding sections of data to be loaded locally and then used will need to be explored.

Logically, the Remote Access Controller is a separate component within the Internet Appliance (Figure 8), but in practice it can easily be integrated into the hardware and software of the Internet Appliance.

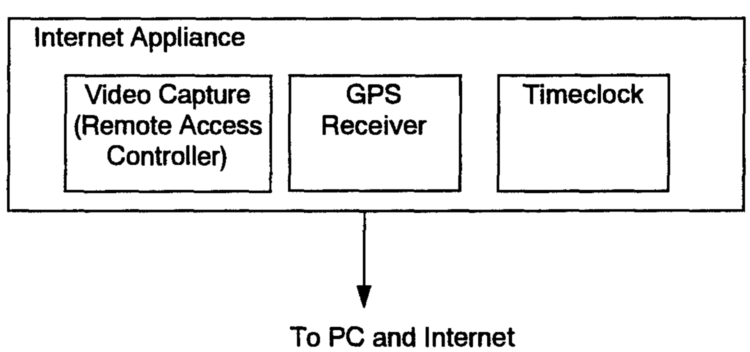


Figure 8: Functional Block Diagram of Internet Appliance

We find that different sensing devices may be used instead of the simple video camera, GPS receiver and timeclock that we have already discussed. These devices may be other kinds of video cameras and GPS receivers, plus entirely different devices such as radar units, infrared sensors, temperature sensors, radiation sensors, etc, and may also be used simultaneously. The PC that drives the Internet Appliance may be tied up with other programs, so it may make sense to offload many of the polling and overhead functions of the sensing devices (i.e. video camera, GPS receiver, radar unit, etc) onto the Internet Appliance's microprocessor. This would keep the PC free to work on other tasks. This may or may not be a firm requirement imposed by the end user, but is a consideration we feel is worth implementing.

An example of the use of the Internet Appliance is for mission rehearsal for the Army. First the Internet Appliance with its video camera, GPS receiver and timeclock is sent out to the actual location where the actual mission is to take place. This may be done by a reconnaissance team, a drone airplane, or perhaps dropped out of an airplane in huge quantities as disposable data collectors. The video camera takes a panoramic video shot of the location, the GPS receiver determines the coordinates and the timeclock's time is recorded. These three pieces of data are uploaded to the internet and made available to those taking place in the mission rehearsal. There may be many or few shots taken, dependent upon the requirements or as the situation dictates.

The users, who may be spread around the globe, but who may be rehearsing for the same mission, can then view the appropriate panoramic scenes. Since GPS and time information is tagged to the visuals, linkage to terrain maps is simple. This information from the Internet Appliance can be tied to other types of data as well.

To make the example even more specific, if a group of Desert Storm troops were planning a maneuver against Bhagdad, drones with Internet Appliances could be flown through the city to get the latest images of the conditions, opposition forces and defenses. These panoramic images (with their corresponding GPS location and time stamps) could be tied into similar panoramic images obtained from spies or from databases built up from tourist shots. These images could be tied into other types of data, for example, engineering and architectural drawings from German construction firms. So when the users rehearsing the mission are immersed in a surround video, they can click on a certain building and see the floorplan. Since this is on the internet, everyone who has a need to know about the mission rehearsal, whether in the Mideast or in Washington can have access to the exact same data at the same time.

It is not necessary that mission rehearsal be limited to activities that will have humans in the area of interest. Some kinds of missions have equipment or ordnance entering the area of interest instead of people. Perhaps this technology can be used to identify targets for bombing or for autonomous vehicles.

Implementation

We have based the Internet Appliance upon a microprocessor controller that provides a wide range of features, including multiple I/O ports, large amount of memory, use of the C-language, a keypad and LCD panel. It is more powerful than the microprocessor we used in the Remote Access Controller. All the features may not be required for a production version, but in the development phase they will provide flexibility and expandability. Thus we will be able to quickly implement changes and perform engineering and user tests.

We have obtained a new microprocessor controller system, assembled and debugged the hardware and are in the process of porting the Remote Access Controller software from the Basic Stamp to this new controller. The new controller is the BirdBox(tm) from Tern Inc which is driven by a C-Engine(tm) microprocessor running the C language. It has four serial ports, twelve digital outputs and 5 digital inputs that can be configured as a parallel port. There are seven channels of comparitor inputs with software-programmable threshold voltages. There are also seven channels of solenoid drivers, two PDC (Portable Data Carrier) ports, a 16x2 character LCD display, an interrupt-driven 3x4 keypad, a reset pushbutton, red and green LEDs, a beeper, and DC power jack, all packaged in a 4.8 x 3.7 x 1.5 inch enclosure.

We have also obtained a Garmin 45 GPS receiver and are using it to familiarize ourselves with GPS technology. This is a popular, low-cost, commercially available unit that is readily available in boating shops and sporting goods stores, and is about the size of a cigarette pack. Although it is small and inexpensive, it is proving to be reasonably accurate, precise and rugged.

The Internet Appliance Design Code

A prototype of the Internet Appliance controller code was implemented using C for portability. This code can be used in most embedded systems on any Intel compatible microprocessor. An example would be the BirdBox C-Engine microprocessor-based controller running the C-language, from Tern Inc.

The BirdBox has a keypad and LCD which will allow for flexibility of design. The software can be modified to accept inputs from the keypad on the fly, while displaying values on the LCD. This will provide a greater range of engineering design and test capabilities, especially when devices of different masses and moments of inertia are to be driven. For example, the controls algorithms can be interactively modified to provide the best tilt and pan response subjectively, without having to devote time to control system design and debugging.

Another problem is that the Engagement Response Server needs to obtain and present the data from different databases. We take a concrete example. Many architects, civil engineers use AutoCAD for their drawings. Each of these users, although they provide the drawings in the same format (dxf, dwg, dwf), use different classification systems for their drawings. Since they may all concern the same virtual reality database, e.g. a virtual scene walkthrough of an enemy structure, we need to correlate and classify them to a common system so that the field commander may move easily from the architect's drawing to the mechanical and piping layouts to the virtual views. This is also necessary because we also envision the system to be used to keep these drawings (or references to these drawings) and the virtual scenes in a database to be used sometimes very much later in time. Unless some classification system is used, it will be hard to retrieve the drawings or data.

To analyze this problem further and determine its feasibility, we developed an example of a classification system that can be used to unify the drawings from architects, civil engineers and mechanical engineers. This is only an example and not meant necessarily to be implemented as it stands.

Commercialization

Relationships

Xtensory has formed working relationships several companies that are involved with avatars, VRML and internet communications. These will enable us to leverage certain technologies for enhancement.

- *Progressive Networks Inc.* We are members of their Developer and Reseller Program, and are working with their RealAudio product to send audio over the internet. In particular, the goals are to embed audio files into mission rehearsal scenes and to broadcast audio over the internet (live or pre-recorded). This is particularly interesting not only because any immersive visualization is empty without the audio component, but also because this technology offers an alternate means of audio communication which is very important on a battlefield.

- *Live Picture Inc*, headed by John Scully (former head of Apple Computer Company, has developed the Imaging for Internet solution, which allows FlashPix images to be sent across the internet at any resolution. In addition, it can deliver FlashPix images directly from the Imaging for Internet server to any standard VRML 2.0 browser, permitting photorealism. It also provides leading image editing and network imaging solutions based on FlashPix, FITS, and internet technologies for professional and home use.

- *Real Space Inc.* has created an extension to VRML 2.0 for adding photographic environments and objects to virtual worlds. This extension allows panoramic images, video and image-based objects to be used in the creation of realistic virtual scenes and avatars. Their RealVR software can be used to created virtual worlds that are open to the addition of IRC chat, audio and video. Their Vistagrapher software makes constructing panoramas (via stitching together individual snapshots) simple and easy.

- *LiveWorld Productions Inc* was founded by key employees from Apple's eWorld internet group. It is an internet content studio that creates and operates original programming with a focus on community, audience participation, time and events.

- *Seismic Entertainment Inc.* (also founded by John Scully) is a creator of virtual exploration experiences. Working with organizations such as NASA and the Smithsonian it used virtual reality technologies to recreate expeditions to Antarctica, Mayan ruins and the surface of Mars. Seismic focuses on providing high quality exploration and adventure content and experiences.

Commercialization

One of our beliefs at Xtensory is to get feedback from real users rather than remaining in isolation in our lab. Our remote control camera as implemented above was demonstrated to Compaq Computer Corp.'s management for potential use on some upcoming projects. We also demonstrated to Heitman Retail Properties, a billion dollar property management company (owners and operators of more than 25 upscale shopping malls nationwide). As a result of this demonstration, discussions are taking place with Heitman to use the video component on three actual projects. Feedback on these project tests will be used to refine the interface to provide a more intuitive and "immersive" visualization capability. In addition, we have obtained contingent Phase II funding commitments from Compaq and Heitman.

Field Tests for Key Design Components

We have always believed in testing and working with real customers for successful eventual commercialization, even at the design stage. This has now resulted in a commitment to use as a testbed three different shopping mall projects, now in the early stages, spread over four states and with fifteen different companies. (One of these is in Sarasota, Florida, which is convenient to get input from the US Army Research Institute Simulator Systems Research Unit on the military requirements.) We propose to design and test various aspects of the engagement resource server using a real test bed over the next few months. We will test aspects of the server on a testbed with:

- Many diverse parties with legacy systems which need to be coordinated. This is a situation which certainly applies in the military arena.

- Across many different states and time zones.

- Mission rehearsal requires integration of visual data with spatial, textual, CAD and terrain data.

This working with an actual user has been very beneficial. Even at this early stage, requirements have surfaced which would have been very difficult to identify on our own. These requirements are:

- We need to be able to integrate various numbering methods from all the legacy systems used by the various parties, e.g. engineers use a different numbering system from project managers. In other words, everyone uses their own numbering system for CAD data, maps, drawings, etc. This is a common problem and one that is shared in a military context. If the Engagement Response Server is to act as a mission rehearsal function, it needs to be able to integrate data from different sources to be understandable to the user.

- The requirement for location data such as GPS data and its integration of visual, GPS, terrain, CAD, map and textual data from remote sites for mission rehearsal.

- A core requirement is coordination of various schedules from different parties.

- Need to surface the key criteria or "success" criteria for the mission rehearsal from the commander or ultimate decision maker to all the many players in the mission.

- Requirement for continuous and up-to-date real-time transfer of drawings (CAD data) and integration of schedules.

- Concrete applications of the Remote Access Controller and its various attachments such as video and digital cameras to provide an immersive mission rehearsal capability for various parties as they tackle various aspects of the entire project.

Appendix A: Review of Unit Performance Summaries

Introduction

In order to get a better understanding of the kind of information that is pertinent for the Engagement Response Server we requested some sample data from the Army. We were provided with a document that included data from actual Army units conducting engagements during simulated combat at the National Training Center. This document is a "Unit Performance Summaries for NTC Rotation" for a certain unit. References to specific units and dates have been blacked out to retain confidentiality, but these deletions do not detract from the overall usefulness of this document for our purposes. This document summarizes unit performance for an Armor Task Force and Field Artillery Battalion, and the sources of this document were the Armor Task Force and Fire Support Take Home Packages.

The Unit Performance Summary document uses tables, bar charts and a brief three-paragraph format to summarize the performance of activities, and occasionally uses timelines where necessary. The presentation of information is very structured and categorized. The three paragraphs summarize the Plan, Prepare and Execute actions. The information that is being reported on can typically be characterized as relating to some measure of human activity or accomplishment, as opposed to easily measured numerical data. We studied the nature of this information and how it can be adequately used.

Analysis of Unit Performance Summary

We isolated several main problem areas.

One of those isolated was coordination. In many places within the document, coordination problems are described. In intelligence under fire support it is mentioned "FSO developed three separate fire support plans" none of which supported the commander's intent. Under Command and control, it is mentioned "commander's guidance did not reach the staff."

The second area we would like to address is scheduling. To quote "once the lead elements of the TF were able to view the objective, they had to wait for the artillery and mortars to catch up before indirect fires could be delivered on the objective." Also "there was no coordinated air defense planning by the battalion for these operations."

The simulated combat exercises described in the document have different goals at different levels in the hierarchy of the Army. Different users have drastically different needs and requirements. Within the browser it is easy to tailor the presentation of HTML pages and links so that different users see different requirements.

The promise of the above approach for the commander is that it can tie together information in a visual and interactive manner. The server using standard "off-the-shelf" hardware and software can provide rich multimedia information. And with the engagement response system he can visualize and experience actions, making the information tangible. It can be easier to evaluate something if you can see events happening, compared to reading a paper summary report. It is definitely a requirement that all the senses be used to enhance the situational awareness..

Another benefit is to use the engagement response system to organize mission planning process. Appropriate forms and checks can be built into the engagement server. Another standard internet technique, the "cookie" mechanism can be used to embed and identify every user of the engagement system enabling the tracking of responses and check lists to see that all forms have been provided.

It is important to provide a framework to allow the users the capability of inputting the information it wants, as well as allowing them the ability to modify it.

Appendix B: The Internet Appliance User's Guide, Version 0.2

The Internet Appliance
User's Guide for Version 0.2

Contract Number: DASW01-96-C-0046

SBIR A95-089

Enhanced Virtual Presence for Immersive Visualization of Complex Situations for Mission Rehearsal

November 3, 1996

Submitted To

Mr. Bob G. Witmer
US Army Research Institute Simulator Systems Research Unit
12350 Research Parkway
Orlando, FL 32826-3276

Tel: 407/384-3995
Fax: 407/384-3999
Email: witmerb@stricom.army.mil

Submitted By

Xtensory Inc
140 Sunridge Drive
Scotts Valley, CA 95066

Tel: 408/439-0600
Fax: 408/439-8845
Email: cutt@netcom.com
URL: http://www.xtensory.com

Introduction

This manual describes how to install and operate Version 0.2 of the Internet Appliance.

This manual is for the engineering prototype of the remote access controller, Version 0.2. This version of the remote access consists of two distinct parts, a video component and a controller component, requiring two separate PCs. See Figure 1.

This system, when fully set up, consists of a the camera/motion platform, a PC that runs the video software (Video PC), and a second PC (Control PC) that controls the platform motions. This separation was done to simplify the tasks and allow the basics to be tackled first using available equipment in our lab. It allows complete de-coupling of tilt/pan platform control functions from video play functions. Future versions of the Remote Access Controller are planned to have the video and control software running on a single PC.

Getting Started

The shipping container contains the following parts:

- Controller Box
- Power Supply and Cable
- Tilt/Pan Platform with Camera
- Connectix QuickCam Software (3 Floppies)
- Connectix QuickCam Manual
- Procomm Modem Software (1 Floppy).

Please check that all the parts arrived safely.

The equipment required to run the remote access video controller are:

PC compatible computer running Windows 3.1 or Windows 95 (Not Windows NT). In this document this will be referred to as the "Video PC."
PC compatible running DOS or Windows 3.1. In this document this will be referred to as the "Control PC."

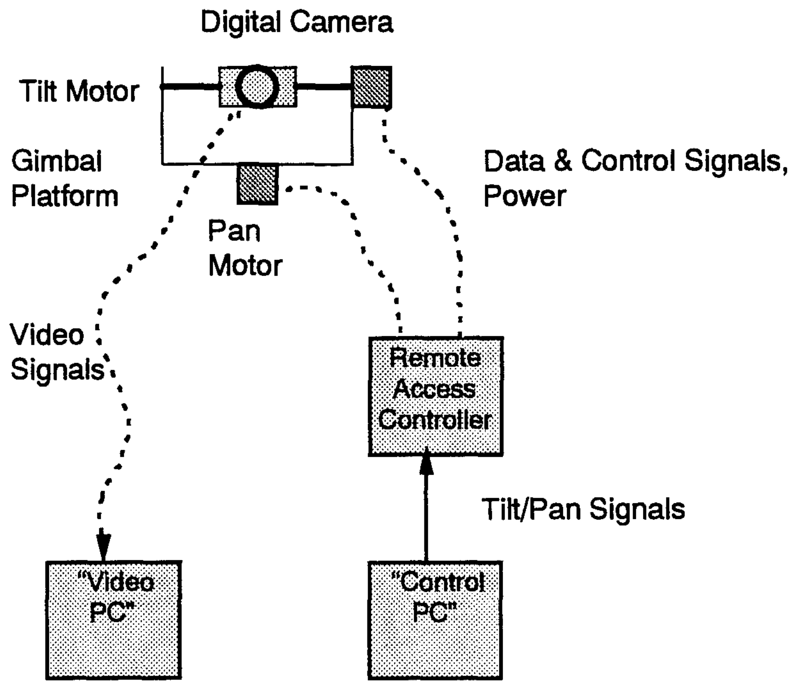


Figure 1: Implementation of Remote Access Video System

Installing the Internet Appliance

Connect the components of the Internet Appliance system to the two PCs as shown in Figure 1. The cables are labeled.

Video PC

Note that the cable from the video camera has two tails, one attaches to the parallel port and the other to the keyboard port (as a "passthru" connector). See the Connectix User's Guide for additional information.

The Video PC requires Windows 3.1 or Windows 95 (Not Windows NT). Install the Connectix software.

Control PC

Plug the controller into a serial port on the Control PC. This signal cable has several 9-pin connectors on it. A 9-pin to 25-pin adapter is provided in case your PC has a 25-pin serial port.

The Control PC requires DOS or Windows 3.1. Install the Procomm software.

Running the Internet Appliance

Video PC

Run the Connectix software. The video window displays whatever the camera happens to be pointing at.

Control PC

Run the Procomm program. The default settings have been set up and do not require setting up.

With the appropriate port setting (Com 1) and baud rate (2400, N, 8, 1) you are now ready to enter keyboard commands to control the motors.

Powering Up the System

Turn on the power supply.

The power supply should be turned off whenever not using this system. This is to prevent overheating the stepper motors.

Controlling the Tilt/Pan Platform

Commands consist of a number and a letter, typed in from the Control PC keyboard. Each command is of the form:

<nnn><space><cmd>

where <nnn> is a decimal number and <cmd> is a single character. The number is the amount of "ticks" the stepper motor will index, each tick being 3.75 degrees (96 steps per revolution). The letters that control motion are l (left), r (right), u (up), d (down). The protocol is to enter a number, a space, and a letter. A "1" is returned to the screen to indicate a correct response by the stepper, and a "0" is returned to indicate a fault.

In addition to the commands to move the steppers, additional commands are available. To change the time in between steps "w" is used. To set a new zero point "s" is used, to move to the zero point "z" is used. To query for ready "q" is used. A zero point can be selected, for example, to view a certain object. Once set, that point can immediately be recalled by using the "z" command.

Microcontroller Hardware and Software

Microcontroller Hardware

We have based the Internet Appliance upon a microprocessor controller that provides a wide range of features, including multiple I/O ports, large amount of memory, use of the C-language, a keypad and LCD panel. It is more powerful than the microprocessor we used in the Remote Access Controller. All the features may not be required for a production version, but in the development phase they will provide flexibility and expandability. Thus we will be able to quickly implement changes and perform engineering and user tests.

The new controller is the BirdBox(tm) from Tern Inc which is driven by a C-Engine(tm) microprocessor running the C language. It has four serial ports, twelve digital outputs and 5 digital inputs that can be configured as a parallel port. There are seven channels of comparitor inputs with software-programmable threshold voltages. There are also seven channels of solenoid drivers, two PDC (Portable Data Carrier) ports, a 16x2 character LCD display, an interrupt-driven 3x4 keypad, a reset pushbutton, red and green LEDs, a beeper, and DC power jack, all packaged in a 4.8 x 3.7 x 1.5 inch enclosure.

Microcontroller Software

Below is a pseudocode segment for the Internet Appliance.

```
REM  This is the gimbal-eye driver.
REM  It controls 2 stepper motors to move the ball camera
REM  up/down (via one stepper) and right/left (via the
REM  other stepper). Each direction is limited, i.e. the yaw
REM  is not 360 degrees.
REM  The serial port IN is the command stream.
REM  Unknown at this point if we are going to use the
REM  serial OUT for ACK
REM  Each command is of the form:        <nnn><space><cmd>
REM  where <nnn> is a decimal number && <cmd> is a
REM  single char. In all cases is this format used, even
REM  though the command may not require (or desire) a
REM  number - eases parsing.
REM  NOTE: The basic stamp parser requires that there
REM  be a char after the number, and that the char is
REM  ignored . Thus, we put in a space to throw away.
REM
REM  The commands are:
REM        l      move left
REM        r      move right
REM        u      move up
```

```
REM        d    move down
REM        w    change wait time (in millisec, 1..255)
REM        s    set zero point
REM        z    move to zero point
REM        q    query ready
REM
REM  NOTE: all commands return a 0 for ACK
REM  The move commands return the ACK BEFORE the
REM  move.
REM  The "q" command is used to determine when the
REM  system is ready.
REM  The zero point is set at the factory (in EEPROM).
REM  The unit will store the current pitch and yaw
REM  positions in EEPROM.
REM  This will allow the unit to "zero" at power-up from
REM  wherever it was.
```

Appendix C: References

Mosteller, F. and Nogee, P. 1951. An experimental measurement of utility. Journal of Political Economy 59, October, 371-404.

Shubik, M. 1962. Some experimental non zero sum games with lack of information about the rules. Management Science 81(2), January, 215-34.

Siegel, S. and Fouraker, L. 1960. Bargaining and Group Decision Making. New York: McGraw-Hill.

Wilson, R. 1984. Multilateral exchange. Working Paper No. 7, Stanford University, August.

Appendix D: Standards - Drawing Classification System

The drawing classification system is described here in greater detail. The key elements of a standardized numbering system for a virtual reality database concerning the drawings portion are:

- Virtual Reality Database Name. Entries no more than 4 characters long, of the format CXXX, where

    C — A character field representing the status of the virtual reality database. Various status fields are:

    - S — Sketch
    - W — Working Drawings
    - C — Construction
    - A — As-built.

    A 3 alphanumeric virtual reality database code. This code field will be released by the virtual reality database owner to all user groups working on the virtual reality database.

- Virtual Reality Database Description. With entries no more than 40 characters long. A brief description of the virtual reality database.
- Virtual Reality Database Identification. With entries no more than 15 characters long. This is needed for accounting purposes. The VR database identification will be generated by the virtual database owner.
- Virtual Reality Database Manager. With entries no more than 14 characters long. The name of the field commander, supervisor or executive in charge of the virtual reality database.
- The application software interfaces (legacy system interfaces) to be used throughout the entire virtual reality database. This will include the related databases (maybe not) that need to be created. Each user group must inform the Engagement Response Server of the need to (or not to) create a database.
- Create the virtual reality database (if applicable), attach the relative schemas (if applicable).
- Assign the virtual reality database and the application software to the operatives. Similar setup of allocating virtual reality database to operatives can be performed on the internet to the family of Engagement Response Servers.

Virtual Reality Database Directories

All virtual reality databases will be accessible via customized interfaces. The interfaces will give user access to all virtual reality databases owned by the same group.

The level of access to all virtual reality databases is controlled by field commanders and above. It is separated into two sections.

```
UNIX End        |        PC End
                |
/usr2/tchaive/VR databases/proj_one
/usr2/tchaive/VR databases/proj_two
                |
```

Here, the field commanders are able to view all virtual reality databases that are owned by the group. They must exercise care in maintaining all virtual reality databases. They must not simply assign virtual reality databases to users.

For operatives:

```
UNIX End        |        PC End
                |
/usr2/tkhoohl/proj_one
                |
```

In the operative level, the user is only able to view one virtual reality database at a time. Note the operative will still have the permissions necessary to remove files.

File Organization

Introduction

This section details several strategies for the organization of CAD drawing files for a given virtual reality database, user group, application software product. The procedures described here are used solely to help all parties concerned identify information. It is recommended that all user groups follow these standard conventions and must not move away from them. Of course the organization is envisioned to be coupled with hyperspace viewing techniques such as those pioneered by SRI (Stanford Research Institute) for navigation in three dimensions of web pages.

Reference Files

Reference links are very important in any classification system. It is envisioned that only the most essential data is kept by the Engagement Response Server. In keeping with the internet web philosophy, dynamic links will be used for example to dynamically link a drawing to one or more other drawings. The drawings themselves may be kept by the responsible organization. Another value in doing this is that information can be distributed to many drawings in a virtual reality database and as the base information changes each of the linked drawings is automatically updated.

The most obvious application of this capability is the Border Sheet associated with a particular virtual reality database dynamic link. Most of the information on each drawing (that is virtual reality database name, border, work order number) is identical. With this in mind, a single border sheet can be created for the whole virtual reality database, referenced to each drawing, and overlaid with the drawing-specific information (that is sheet number, title, etc). If the virtual reality database title needs to change, the change only needs to be made to the original and it will automatically show up on all drawings to which it has been referenced. A similar approach may be used to incorporate parts of a design file into many different drawings with the same degree of spontaneous update.

Note: The "/central/revision" and "/central/standards" directories are created for this purpose.

Filename Format

The following convention has been adopted and must be followed by all user groups. Failing to do so may have adverse effects when referencing a drawing.

They are

xxx xxxx.cc

where

x Group Name

Group	Valid Group Name
Field Command	1
Civil Engineering	c
Structural Engineering	d
Mechanical Engineering	m
Electrical Engineering	r

the next

xxx Virtual reality database Code, a 3 digit alphanumeric code, to be determined for a particular virtual reality database.

the next

xxxx Drawing Numbers, a 4 alphanumeric user definable field. All drawing numbers which are less than 4 characters long are to be padded with leading zeroes.

the next

cc File Type

The following list has been adopted and must be followed by all user groups. This is to ensure that there will be no filename clashes between groups.

Description	File Type
Architectural	
Detail	de

Elevation	el
Plan	pl
Section	se
Site Layout	sl
Ceiling	ce
Furniture Layout	fl
3D Modeling	mo
Electrical	
Detail	de
Lighting Layout & Fans	ll
Power Layout & Accessories	pl
Schematic	sh
Single Line	sl
Communication & Computers	cm
Fire Alarm	fe
Building & Energy Management	bm
Lightning Protection & Earth	le
Outdoor Lighting	ol
Internal Cabling	ic
External Cabling	ex
High Tension	ht
CD Installation	cd
Generators	gn
Uninterruptable Power Supply	us
Audio & Video Systems	av
Security System	sy
Mechanical	
Air Conditioning, Mechanical, Ventilation	ac
Fire-related Systems	fs
Plumbing Water & Gas	pl
Lifts, Escalators, Dumbwaiter	li
Medical Gas	mg
Irrigation & Foundation	ir
Kitchen Equipment	ke
Laundry Equipment	le
Sanitary-related Equipment	a
Civil	
Elevation	el
Plan	pl
Section	se
Survey	sv
Survey Sheet Layout	sp
Design Sheet Layout	sd
Existing Sideroads/Buildings	rb
Existing Trees	et

Topographic Survey	pt
Cadastral Survey	cs
Setting Out/Key Plan	sl
Traffic Plan/Site Plan	pn
Longitudinal Section	ls
Detail Plan	de
Structural	
Column Loading	cl
Floor Plan	pl
Column Schedule	cs
General Notes	gn
Drawing List	dl
Typical Details	td
Miscellaneous Details	md
Beam Details	
1st Story	1b
2nd Story	2b etc
Slab Details	
1st Story	1b
2nd Story	bc
Sub-station	ss
Roof Truss	rt

c Revision Number, a single character field. Files that are first revised will have a revision number of "a." Subsequent revisions will be "b" and "c" etc. Files with no revision will have a null entry.

Enhancement: Terrain and Maps

The classification system above can also be extended to include identification of appropriate terrain and area maps. It can be used to manage the placement of overlays on the maps and provide a global map structure to point to other information for mission rehearsal. One area which of course is necessary is to cater to different coordinate systems. See "Technical Reference Datums, Projections, Grids and Common Coordinate Systems, Transformation of", Department of Defense, MIL-HDBK_600008, May 1991 DMASC-SGS, 8613 Lee Highway, Fairfax, Virginia 22031-2137.