PERGAMON

Computers & Graphics 26 (2002) 535–555

COMPUTERS & GRAPHICS

www.elsevier.com/locate/cag

Experiments in Immersive Virtual Reality for Scientific Visualization

Andries van Dam, David H. Laidlaw, Rosemary Michelle Simpson*

Department of Computer Science, Brown University, 115 Waterman Street, Providence, RI 02906, USA

Abstract

This article provides a snapshot of immersive virtual reality (IVR) use for scientific visualization, in the context of the evolution of computing in general and of user interfaces in particular. The main thesis of this article is that IVR has great potential for dealing with the serious problem of exponentially growing scientific datasets. Our ability to produce large datasets both through numerical simulation and through data acquisition via sensors is outrunning our ability to make sense of those datasets. While our idea of "large" datasets used to be measured in hundreds of gigabytes, based at least in part on what we could easily store, manipulate, and display in real time, today's science and engineering are producing terabytes and soon even petabytes, both from observation via sensors and as output from numerical simulation. Clearly, visualization by itself will not solve the problem of understanding truly large datasets that would overwhelm both display capacity and the human visual system. We advocate a human–computer partnership that draws on the strengths of each partner, with algorithmic culling and feature-detection used to identify the small fraction of the data that should be visually examined in detail by the human. Our hope is that IVR will be a potent tool to let humans "see" patterns, trends, and anomalies in their data well beyond what they can do with conventional 3D desktop displays. © 2002 Published by Elsevier Science Ltd.

1. Overview

This article provides a snapshot of the use of immersive virtual reality (IVR) for scientific visualization, in the context of the evolution of computing in general and of user interfaces in particular. By IVR we mean a compelling psychophysical illusion of an interactive 3D environment surrounding the user that is characterized by presence: the feeling of "being there", a sense of the reality of both objects and the user in the scene. The illusion is produced by hardware, software, and interaction devices that provide head-tracked, wide-field-of-view stereo to present a human-centric point of view; typical hardware includes head-mounted displays and CAVE™s[1] and their derivatives [1]. Spatial sound and haptic (force and touch) "display" devices are often used to add realism to the visual illusion. IVR can be seen as an extension of desktop computing in much the same way that an IMAX stereo experience can be seen as an extension of watching the same content on a standard television set. IVR, however, can provide an even more compelling experience than IMAX stereo because it is not only immersive but can be fully interactive. Indeed, it is a fine example of multimodal, post-WIMP[2] interfaces in that it emphasizes the natural use of multiple sensory channels in parallel: it mimics the way in which people interact with the real physical world, using as much as possible "whole-body interaction".

The main thesis of this article and its predecessor [2] is that IVR has great potential for dealing with the serious problem of exponentially growing scientific datasets. Our ability to produce large datasets, both through

*Corresponding author.

E-mail address: rms@cs.brown.edu (R.M. Simpson).

[1] We use the term Cave to denote both the original CAVE developed at the University of Illinois' Electronic Visualization Laboratory and CAVE-style derivatives.

[2] WIMP interface denotes Windows, Icons, Menus, and Pointing, the standard desktop graphical user interface designed at Xerox PARC in the early 1970s.

PII: S0097-8493(02)00113-9

numerical simulation and through data acquisition via sensors, is outrunning our ability to make sense of them. While "large" datasets used to be measured in hundreds of gigabytes, based at least in part on what we could easily store, manipulate, and display in real time, today's science and engineering produce terabytes and soon even petabytes. Clearly, visualization by itself will not solve the problem of understanding truly large datasets—that would overwhelm both display capacity and the human visual system. We advocate a human–computer partnership that draws on the strengths of each partner and uses algorithmic culling and feature detection to identify the small fraction of the data that should be examined in detail by humans. Our hope is that IVR will be a potent tool to let humans "see" patterns, trends, and anomalies in their data well beyond what conventional 3D desktop displays can provide.

Below we first set context by summarizing some important trends in computing and communication in general and in user interfaces in particular. We then discuss the use of IVR in scientific visualization, illustrated primarily by ongoing work at Brown University, and then briefly describe tele-immersion work being done in conjunction with researchers at the University of North Carolina/Chapel Hill and the University of Pennsylvania. We conclude by summarizing some interesting research issues in IVR and its use in scientific visualization.

Since space limitations prevent a comprehensive treatment, we focus on themes especially relevant to our graphics work at Brown University, with the hope that they will be of interest to a broader community. The paper builds on [2,3] and upon ideas addressed by the final author in his keynote address for Jose Encarnacao's 60th birthday festival in Darmstadt, Germany in May 2001.

2. Trends in computing and communication

2.1. Price/performance

In the early 1980s Raj Reddy and his colleagues at CMU coined the term "3M Machine" [4] to characterize personal workstations with an MIP of performance, a megabyte of memory, and a megapixel display (as well as a megabit network connection). Since then Moore's "law" for ICs, clever electromechanical device engineering, and the commoditization of all personal computer components have led to tremendous progress in hardware price/performance, of roughly three to four orders of magnitude in operating speed and memory size (both main and disk), as well as in graphics performance. Furthermore, today's commodity 3G PC with incredible real-time 3D graphics costs a tiny fraction of the 3M workstation. This continuing technology improvement, unprecedented in human experience, drives all of the other trends we describe below.

However, despite these spectacular price/performance improvements, desktop productivity applications and their user interfaces are still for the most part stuck in the rut of WIMP GUIs, whose commercial use dates from the early 1980s. These tools do not run noticeably (i.e., orders of magnitude) faster than in the 1980s and are not substantially easier to use. We are still looking for breakthroughs in the human–computer interface, both for the familiar desktop computer and for the collections of newer platforms at our disposal, ranging from PDAs, tablet PCs, and projection displays to immersive VR and augmented reality (AR) environments.

A common question is whether we can make good use of, indeed really need, ever-increasing compute power? The answer is a vigorous 'yes', for two reasons: (1) even for ordinary desktop computing we need the power to provide a much more satisfactory user experience through richer, yet easier-to-use interfaces, a quest that Raj Reddy articulated in his call for SILK interfaces (speech-, image-, language-, and knowledge-based) [5], and (2) we are still orders of magnitude away from having enough capacity to handle realistically large datasets in immersive environments, as we will explain below.

2.2. Ubiquitous and mobile computing

2.2.1. Wireless

While the wireless community suffered the same fate as the rest of high tech during 2001, the perceived need for it grows stronger each day, as shown by ACMs granting of permanent SIG status to SigMobile. Neighborhood area networks (NANs), also called WiFi are springing up, based on IEEE 802.11b standards, as part of a "communication for the people" grassroots movement. Wireless technology is also increasingly being used for peripheral devices, e.g., the Comdex 2001 Best of Show award to Senseboard Technologies' virtual keyboard, a wireless hand device that enables PDA and wearable computer users to turn any surface into a virtual keyboard. Bill Gates' keynote at CES 2002 showed that even the desktop is becoming wireless—he displayed ViewSonic's AirPanel 150, a wireless monitor for use within a LAN-type environment, e.g., home or office building.

2.2.2. Profusion of form factors

The CES 2002 show was striking in its shift from analog to digital devices. The prediction [3] that "All the digital accessories cluttering our briefcases, pocketbooks, and belts will merge into much more general purpose communications devices with knowledge of our preferences, our context, and our geographic location."

has started to come true, as exemplified by Danger, Inc.'s announcement of the multifunction Hiptop combination PDA, telephone, Internet browser, email and instant messaging utility, and camera.

The introduction at MacWorld in January 2002 of the new iMac form factor further emphasizes the move of the industry away from desktop-hogging boxes and large-footprint CRT monitors to flat screen and small footprint devices. We will see growth at the other end of the display spectrum with larger LCD and plasma panels, Digital Light Processing™ projection displays [6] for office environments, and the tantalizing prospect of flexible and arbitrary-sized geometry provided by light-emitting polymers [7], which in early 2002 were still low resolution and monochrome but showed great promise for medium- to high-resolution color displays.

2.2.3. Embedded and invisible devices

Another way in which computing will diversify is to become embedded (along with appropriate sensors and actuators) in our appliances, our rooms and furniture ("smart environments"), our vehicles, our clothing, and even in our bodies. The critical core technologies speeding this trend along are nanotechnology in general and MEMS in particular. Most of these devices will either have no interfaces at all or make use of very simple task-oriented interfaces. This will be especially prevalent in monitoring and regulating implants in our bodies (e.g., insulin units), and is already present in the dozens of "computers" embedded in our automobiles.

2.2.4. Federation of devices

Fig. 1 shows a future computing environment consisting of potentially hundreds of devices per user and millions of interconnected users. The devices will range from conventional desktop computers to IVR environments with their head-tracked stereo glasses, data gloves, and speech recognition. The dispersal of computing power into a multitude of devices in multiple form factors requires new coordinated approaches to software design and data interchange protocols [8] if serious data balkanization (the so-called "data silos" problem) is to be avoided.

Many of these devices will be directed by autonomous agents rather than explicit user control. All will need to synchronize data, and data flow must be seamless as users move from one location to another, from one device to another. At the same time, approaches to solving this problem, which typically involve location- and identity-sensing technology, must also address privacy issues, already a serious problem with both Internet access and proposed security legislation.

2.3. "Netification" of middleware and applications

In parallel with physical dispersal of devices is the logical dispersal of computing, which generalizes the decades-old client–server model promulgated by PARC as part of the workstation environment that also brought us the WIMP interface. The key idea here is that monolithic applications that today are at most

Fig. 1. The computing environment of the future will use the standard networking 'cloud' to make possible both formal and informal collaboration of users through a variety of post-WIMP techniques. Image courtesy of Anne Morgan Spalter and Sittha Sukkasi.

distributed across a server and its clients will be deconstructed into more or less fine-grained components that exist in the network as a set of services to be assembled and integrated on demand. It is not obvious that such large applications as Microsoft Word or an Oracle relational database will be so deconstructed, but clearly we need to master the art of truly distributed programming in order to control federations of devices. Contending application frameworks for network services include Microsoft's. Net [9], IBM's WebSphere [10], and Sun Microsystem's Sun ONE [11].

2.4. IVR and AR

While IVR and AR (Augmented Reality, the blending of computer-generated information and the real world in one display) [12] show progress and IVR in particular is used in production by high-end design companies and by research labs, they are not yet mainstream because serious technological problems persist and their price/performance has not attained the level required for commoditization. They remain niche technologies, though this may be changing as the costs and hazards of travel combined with accelerating globalization drive one aggressive use of this technology, tele-immersion (see Section 6).

While XEROX PARC's Mark Weiser viewed ubiquitous computing as the antithesis of virtual reality, in fact we are beginning to see these two technologies start to blend in two major areas: the use of PDAs and other handheld devices within IVR environments, and tele-immersion, which aims to unify the working space of geographically separated but network-connected offices.

2.5. Collaboration as major working and playing mode

People have always worked in formal and informal groups and since the advent of the computer have wanted computer support for this activity—the ACM CSCW conferences began in 1986 and have their roots in 20 years' prior work on computers and the organization of work [13]. Until recently the groupware research community was fairly distinct from the rest of the human–computer interface community as represented by SIGCHI, but we are increasingly seeing intellectual boundaries dissolve as the increasing ubiquity and dispersion of computing resources blurs the lines between the tasks being executed at a given time, where they are done (home, office, en route), and who does them.

The notion of a single user doing a single task on a single computer has given way to team-based, highly multitasked computer-assisted activities. For example, in the spirit of Weiser's vision, Bill Buxton, Alias/Wavefront's chief scientist, paints a vivid picture of impromptu meetings facilitated by office and corridor walls that are interactive display surfaces able to access remote environments and their occupants as needed. While the term collaborative virtual environment (CVE) includes almost any form of computing environment that facilitates collaboration, ranging from desktop-based chat rooms, MOOs, and massive multiplayer games to PowerWalls [14], a high-resolution tiled display, and full room-sized IVR [15], here we discuss only the specific form of CVE called tele-immersion. The National Tele-Immersion Initiative (NTII) project [16], sponsored by Advanced Network & Services, Inc. described it as an advanced form of IVR: "Tele-Immersion will enable users at geographically distributed sites to collaborate in real time in a shared, simulated environment as if they were in the same physical room."

3. Trends in user interfaces

3.1. Introduction

The user interface is the means by which the user controls a computing environment. On the one hand, it should be as simple and transparent as possible, putting the minimal cognitive load on the user by pushing most of the mechanics to the perceptual level to achieve automaticity (much like driving a car). On the other, it must be sufficiently powerful to accomplish its task efficiently. There are well-known tensions between ease of learning and ease of use,[3] between what is best for the novice or occasional user and what is best for the power user, and between a single general-purpose interface style (e.g., WIMP) and multiple specialized ones. Indeed, the case Bill Buxton [17] makes for specialized interfaces, is compelling although there will continue to be utility in general-purpose devices and interfaces.

In the limit, we will always have significant user interfaces? One point of view expressed by the first author in [3] is: 'I believe that user–computer interfaces are a necessary evil. They imply a separation between what I want the computer to do and the act of doing it. Ultimately I prefer the computing environment would simply carry out my wishes via a form of telepathy—cogito ergo fac ('I think therefore do it'). Needless to say, such a means of instructing would present unimaginably difficult problems, not just in the complete interpretation and execution of DWIM—do what I mean—but equally in preventing the analog to the Midas touch, where my thoughts are taken too literally and at times when I don't want them acted on at all. 'Be

[3] As easy to learn as WIMP interfaces are supposed to be, it still is distressing to see how much time a user can spend puzzling out what the interface is doing or using trial and error to get it do something that should obviously be easy to do.

careful what you wish for,' indeed…." The cross-disciplinary work essential to developing genuine brain-controlled interaction, however, e.g., work in implanted-electrode control of computer cursors [18], is well underway.

Another view of interfaces does not deny the utility of an omniscient, "virtual Jeeves" who knows one's interests, context, and style and can therefore understand and indeed anticipate one's wishes, but points out that a good tool (e.g., handsaw, tennis racket, violin, and some software) has a good feel (interface) which produces delight in its conscious use. Finally, we certainly need interfaces that operate at a higher level of abstraction, that are more geared to expressing our intent and do not insist on literal specification of each atomic component of an interaction.

3.2. Evolution towards post-WIMP UIs

Shifts in computing environments make possible and mandate shifts in user interfaces. The watershed computing paradigm shift was from batch computing to personal computing, while in personal computing we had a user-interface paradigm shift from textual command lines on character terminals to WIMP GUIs based on bit-mapped raster graphics workstations.

A watershed computing shift comparable in its impact to the move from batch to PCs is now beginning as we evolve from a computing environment that serves a collection of individual users at their desktops communicating loosely via email, chat, or instant messaging to a much more highly distributed environment that serves not only individuals but increasingly, teams of collaborating users, each controlling not just a PC but a federation of devices. This change in computing carries profound user-interface implications. On pocket PCs and PDAs, we see hybrids of WIMP and gestural interfaces; on more embedded devices, e.g., car radios, we see interfaces that directly reflect their analog counterparts; and many devices have no interface at all since we control them indirectly, as when we drive any late-model automobile.

The user-interface needs for handheld devices with small amounts of screen real estate are arguably very different from those for conventional PCs, which are yet again different from those for a wall-sized display. Web design, in particular, must take device characteristics into account [19]. There is, of course, a tension between the desire to create interfaces that take maximum advantage of a device's physical form factor and capabilities, and the desire to have the operating environment and applications be largely independent of such device characteristics. Pocket PCs today typically run subsets of the operating environment and of the applications, and the limitations of these reduced versions do intrude.

The handheld device that contains addresses and calendars in one context may connect to the Internet to search for a nearby restaurant in another, and in yet another context is used in a design meeting to control and annotate a design sketch on a large PowerWall display; the same device in the first context is a single-user restricted function device, in the second is a global network front-end, and in the third is an integral part of a multidevice collaborative environment. The user interface changes to reflect not only the application but also the physical environment. For example, PDAs used in a Cave environment must take into account the low light levels and the fact that the user is wearing shutter glasses–both affect the user's ability to read the PDA. However, some capabilities like using pen-based menus for selecting from lists or pen-based interactions for sketching 2D shapes should be feasible in the Cave.

3.3. Post-WIMP UIs for IVR

In IVR environments we see a wide range of user interface experimentation, ranging from work that integrates gestural sketching interfaces into existing applications [20] to bold, indeed sometimes bizarre, experimentation with multi-modal interfaces that appeal in principle to the entire human sensorium. These interfaces are characterized by simultaneous use of multiple sensory channels and whole-body interaction. Users continuously move their bodies and especially their heads as they examine the scene, gesture using various interaction devices such as data gloves, wands, and use PDAs, voice commands, etc. Much decoding of the multiple simultaneous and often high-bandwidth input streams must be done probabilistically, where possible using unification algorithms that support mutual disambiguation between information coming from separate input streams. These media-stream unification algorithms, originally developed for 2D work [21], are now being extended to wireless mobile 3D computing. This powerful and complex post-WIWP non-deterministic interaction style contrasts vividly with the far more limited, sequential and deterministic style of WIMP GUIs restricted to keyboarding and mousing.

3.4. Group interaction

Among the many innovations[4] Doug Engelbart showed in his landmark 1968 demonstration of NLS [22] was live telecollaboration with shared floor control with a collaborator at a remote site. As with his

[4] Hypertext, structured text editing, the mouse and chord keyboard, two-handed interaction, and windows, to name a few.

invention of mouse, it has taken the computing world a very long time to extend his pioneering work, but current applications include collaborative network editing tools such as Microsoft's NetMeeting [23] and the Nebraska application of the Squeak open software project [24], shared whiteboard projects from GMD [25] and Alias/Wavefront's PDA floor-control project for the Portfolio Wall [26].

In a lecture given at Brown in the fall of 2001 Bill Buxton made an interesting observation on the social impact of technological change as seen in the dynamics of an auto design studio. Prior to the development of CAD/CAM software, he says managers could walk around the studio to see what work was underway—the social protocol was collaborative, shared, and open. Removal of the design process to the computer workstation eliminated that social protocol and made it far more difficult to know how designs were progressing. The solution, which kept the advantages of workstation CAD while restoring the shared, collaborative and open aspect, was to create a 50″ plasma display with a touch screen on which working drawings were displayed in public space.

At the bleeding edge of IVR and haptic research, Roger Hubbold of the University of Manchester, UK, in work at UNC Chapel Hill has been developing a simulation of a collaborative task in a shared virtual environment in which two users carry a shared object (a stretcher) in a complex chemical plant. The implementation includes a haptic interface for each user, with the Argonne arm at one end with full-sized surrogate stretcher handles attached, and a desktop Phantom at the other (with minature handles attached), so that forces transmitted through the stretcher from one user to the other can be experienced [27]. In other work, UPenn has been collaborating on work with Drexel and universities in Brazil to develop visual and haptic approaches to collaborative tele-presence in medical robotics applications [28].

3.5. Agent intermediation

Agents are currently used primarily for e-commerce, e.g., purchasing and sales agents, and to some extent in multi-player games. Agents with social interfaces (avatars) act either autonomously or under more direct user control for conversational interaction. In addition to this type of agent, avatars are also used as standins for live human beings, typically simplified, to create a sense of presence of the remote collaborator. In the future, agents may also play a critical support role in IVR scientific visualization applications by pruning and selecting relevant portions of massive datasets for presentation in the IVR environment and possibly even for pointing out features of potential interest.

4. IVR for scientific visualization

4.1. Introduction

4.1.1. Vision: the holodeck

The goal of IVR is to create a realistic (through not necessarily real) world that leverages as much as possible of the full human sensorium for both input and output. This goal resembles descriptions of StarTrek's holodeck, in that IVR creates a virtual environment that is experienced as real and immersive in the sense that you are IN the scene, fully interacting with the objects it contains and vice versa. In the most extreme version it is indistinguishable from reality; if there is a real lion, it can eat you, as in Ray Bradbury's 1950 dystopian story "The Veldt" [29].[5] In 1965 Ivan Sutherland proposed a computer-generated virtual world in which the experience was so real that a bullet would kill [30]. Three years later he had created an HMD that produced a stereo 3D view of a computer-generated room [31], a vital first step towards the ultimate vision.

4.2. Motivation for IVR

4.2.1. Context and presence

Both the feeling of immersion and the sense of presence are enhanced by a substantially wider field of view than is available on a desktop display, and both leverage peripheral vision when working with 3D information. This helps provide situational awareness and context, aids spatial judgments, and enhances navigation and locomotion. The presentation may be further enhanced by aural rendering of spatial 3D sound and by haptic rendering to represent geometries and surface material properties.

UNC, in its Visual Pit project [32] has been conducting a striking series of experiments to measure the effect and amount of presence in an IVR environment, in a project named the Visual Pit. This environment, a testbed for measuring the effects of a stress-inducing virtual environment, contains a wooden catwalk border around an apparent cutout in the floor over a 20′ drop to a room below (see Figs. 2 and 3).

Subjects try to walk across the "open space" while their heart rate, skin conductance, and skin temperature are measured. The illusion is so realistic that many subjects (including the first author) cannot force themselves to step off the catwalk, even though they can feel, by tapping a toe on the floor, that the cutout is in fact solid. The goal is to discover physiological measures of presence that are reliable, valid, sensitive, and objective. In the experiments conducted so far, heart

[5] The first author read this as a young boy 50 years ago, never realizing that he would some day be involved in realizing the technology, if not its distopian use.

Fig. 2. Participant in UNC's Visual Pit study of the effect and amount of presence in a stress-inducing IVR environment.

Fig. 3. The Visual Pit IVR virtual environment.

rate has proven the best measure of an effective sense of presence. Continuing issues include such questions as: What contributes to the experience—have you measured what you think you have measured? What is impact of your inability to see your own body?

4.2.2. 3D spatial relations—body-centred judgements

The immersive surrounding context provides a kinesthetic depth perception that lets users better apprehend 3D structures and spatial relationships. It makes size, distance, and angle judgments easier since it is more like in being in the real world than looking through the screen of a desktop monitor to the world behind it; the advantages arise from the difference between "looking at" a 2D image of a 3D environment on a conventional display screen and "being in" that 3D environment and basing spatial judgments relative to one's own moving body. Looking at a picture of the Grand Canyon, however large, differs fundamentally from being there.

4.2.3. Multidimensional data

Much scientific work involves understanding multi-valued data. For example, the relationships among velocity, vorticity, pressure, and temperature throughout a 3D flow might be critical in understanding a fluid flow problem. Correlations among the many values in such data are often best discovered through human exploration because of our visual system's expertise in finding visual patterns and anomalies. However, representing many values simultaneously is difficult in 2D and even harder in 3D. Experience from art and perceptual psychology has the potential to inspire new, more effective, visual representations for this challenge. Artists over the centuries have evolved a tradition of techniques to create visual representations for particular communication goals. Painting, drawing, graphic design, and sculpture techniques all have potential applicability. The 2D painting-motivated examples in David Laidlaw's work [33] have been extended to show multivalued data on surfaces in 3D. We elaborate on these themes in Section 5.

4.3. Uses of IVR

4.3.1. Human scale

The first and still the most prevalent uses of IVR were in areas that reflected human-scale environments. Such environments are effective because the 1:1 scale builds on lifelong familiarity and aids understanding. Applications include entertainment [34], vehicle design models [35,36], architectural walkthroughs, machinery accessibility and repair studies, and role-playing, especially simulations of vehicles, peacekeeping and rescue operations, and battlefields.[6]

4.3.2. Non-human-scale scientific visualization

From a user interface perspective, perhaps the greatest challenges and rewards lie in the use of IVR for non-human-scale scientific visualization. The need to handle enormous quantities of data, visualize abstract relationships, display size ranges from atomic to cosmic, slow down/speed up/stop multidimensional data flows, or interact with inaccessible areas such as the interior of an artery provides challenges not only for visualization but also for the interaction devices and mechanisms. Medical applications range from psychiatry [37] and pain management [38,39] to non-invasive investigation and "manipulation" of human organs such as the brain and the colon, e.g., surgical preplanning [40] and virtual colonoscopy [41].

4.3.3. Teaching difficult skills

Many skills, such as surgery and nuclear reactor and space shuttle repair, are difficult to teach because they require highly specialized prior training or because the areas involved are inaccessible, dangerous, or fragile. IVR provides a possible solution for this dilemma. Section 6 discusses a current research project on immersive electronic books for teaching surgical procedures. Other educational areas in which IVR could be both useful and cost-effective are seldom needed but critical skills, such as bomb defusing or space shuttle operations.

4.3.4. Tele-immersive collaboration

Since 9/11 the trend towards videoconferencing has accelerated significantly [42], heightening both the expectations and frustrations attendant on a medium that promises distance visual collaboration while denying the simplest real-presence experiences of eye contact and relative position placement [43]. Tele-immersion, which combines IVR with some form of synthetic presence of remote participants, has as its goal allowing physically separated people to interact as they would were they actually in the same physical location. Research projects are underway in several labs [44–46], and have progressed to industrial deployment in the case of VisualEyes [47]. In Section 6, we discuss some aspects of the ongoing research project discussed in [43].

[6] In a conversation with Fred Brooks he stated that his investigations indicate that world wide over 500 Caves are currently in use, the majority in production use for various kinds of industrial design such as vehicle and drug design.

5. Where are we now?

At Brown we have been exploring a number of different IVR applications to study UI research issues in IVR. Our strategy here is in the spirit of Fred Brooks' toolsmithing: we are collaborating with domain experts in several fields to develop applications that address driving problems from other scientific disciplines [48,49]. While the extra effort associated with these collaborations goes beyond just the application development, we believe that they will help us to factor out common patterns from the problems in various disciplines to develop IVR interaction metaphors and visualization techniques that can be generalized. The collaborations also let the domain experts validate new techniques and ensure that they are responsive to the needs of real users.

Our scientific application areas include archaeological data analysis, biological fluid flow (bioflow) visualization, brain white-matter analysis, and Mars terrain exploration. In interacting with scientists in each of these application areas we have made a number of observations about working in IVR, about the effectiveness of visualization methods and user interface techniques for different purposes, and about how users react to IVR. In this section, we summarize each application and discuss related research issues.

5.1. Methodology issues

More than in desktop displays, the visual characteristics of a virtual world are coupled with the user interface. As mentioned above, much of our work combines experience from art and perceptual psychology for inspiration in designing, realizing, and evaluating these visual presentations. We draw inspiration from the artistic techniques to support our communication goals; art history, pedagogy, and methodology, together with art itself, provide both source material and a language for understanding that knowledge. Perceptual psychologists have also developed a body of knowledge about how the human visual system processes visual input; these perception lessons aid in designing the visual representations of IVR user interfaces.

Beyond visual inspiration, perceptual psychology also brings an evaluation methodology to bear on scientific visualization problems. Evaluating visualization methods is difficult because not only are the goals difficult to define, tests that evaluate them meaningfully are difficult to design. Similar issues arise with the methods of evaluating how the human perceptual system works. In essence, we are posing hypotheses about the efficacy of user interfaces and visual representations and testing those hypotheses using human subjects, an experimental process that perceptual psychologists have been developing for decades. Perceptual psychologists, in close collaboration with domain experts and artists, are helping us develop a methodology for evaluating visualization methods. In addition, we will be addressing issues of usability engineering [50] as we proceed.

5.1.1. Design

Both visual design and interaction design play together to create a virtual environment. Virtual environment design, however, goes beyond the typical domain or training of a single designer. Many of the design issues are similar to those of varied areas of design from which we can learn and draw inspiration. Components of architectural and landscape design are relevant for creating and organizing virtual spaces, sculpture and industrial design for finer-grained 3D parts of the environment, illustration for the visual representation of scientific data and the surrounding environment; traditional user-interface design is applicable to some parts of the interaction design, and animation to other parts. The design process is different for all these types of designers, and getting all of the pieces to play together effectively is a challenge, particularly with the constraining needs of the scientific applications. We return to the need to develop a design discipline for our field in the Conclusions.

5.2. Archeological data analysis

Archave, created by Eileen Vote and Daniel Acevedo, was one of Brown's first Cave applications. It was developed in cooperation with Martha Sharp Joukowsky of Brown's Centre for of Old World Archaeology and Art and Brown's NSF-funded SHAPE lab [51], which was set up to develop mathematical and computational tools for use in archaeology. Archave provides virtual access to the excavated finds from the Great Temple site at Petra, Jordon.

Because of the inherent three-dimensionality of the archaeological record, archaeological analysis is a natural application for IVR. Archaeologists analyze excavated finds to understand how a site was used through history and, thus, to understand the peoples and cultures of the past. Digs are documented and published via the excavation team's records: typically dusty notebooks full of handwritten notes, surveys of trench relationships, maps, diagrams, and sometimes a computer database of artifacts transcribed from the notebooks. The spatial relationships among the artifacts are often crucial in their analysis but are very difficult to record. Archave provides an IVR interface that presents the original spatial relationships among the artifacts within a model of the site and the present-day ruins.

The application and use of Archave's IVR interface has been reported elsewhere [2,52,53]. Here we examine some of the user interface and design issues that became apparent during the system's design, development, and testing.

5.2.1. Research issues

Our application design had archaeologists work in the Cave for multi-hour sessions on analysis tasks that would have been quite difficult using traditional analysis tools. They were able to support existing hypotheses about the site as well as to find new insights through newly discovered relationships among the over 250,000 catalogued finds. This experience, while successful, raised many research issues of user interface and visual design, many of which involve taking the Cave experience beyond the demo stage and making it truly useful for archaeologists.

Multivalued visualization: In this archaeological application, each of the many thousands of artifacts is described by dozens of attributes, such as Munsell color, date, historical period, condition, shape, material, and location. Furthermore, the relationships among the artifacts and with the trenches from which they are excavated also must be represented in an easily distinguishable way.

Many workers have tackled this problem outside the Cave, mostly in 2D [33,54–57]. As we move into three dimensions and IVR, the challenges increase. Visual cues that are effective in 2D, such as lighting and shading, in 3D provide the viewer with important shape information. When used to encode data, however, they can be confusing because viewers can interpret data changes as shape changes. As with all design issues, the driving force behind the visual representation is the task to be performed with it. For the relatively simple demo-like evaluations performed so far, only a few attributes of the data are displayed—those essential to a specific task. As the tasks become more exploratory, the visual mapping will need to become more complete and simple designs will no longer suffice. Here is where we hope to exploit perceptual psychology and art history for inspiration as well as studying the design process itself.

Design: Design issues are omnipresent in our Cave applications and several from Archave deserve mention. First, the visual context of the data representation, as well as the visual representation for artifacts themselves, are important in interpreting the data. Initially, we used realistic representations for artifacts within a plausible reconstruction of the temple (see Fig. 4). We found, however, that a more primitive iconic representation of the artifacts, with important data components represented by shape, size, and color, was easier to analyze (see Fig. 5). Our Archave users, who had typically been involved in the excavation process, were also distracted by the reconstruction, since they were far more familiar with the present-day ruins. They were able to work more effectively with a model of the present-day ruins as context in which artifact concentrations are shown as simple 3D icons in saturated colors among a muted view of the present day ruins and transparent trench boundaries.

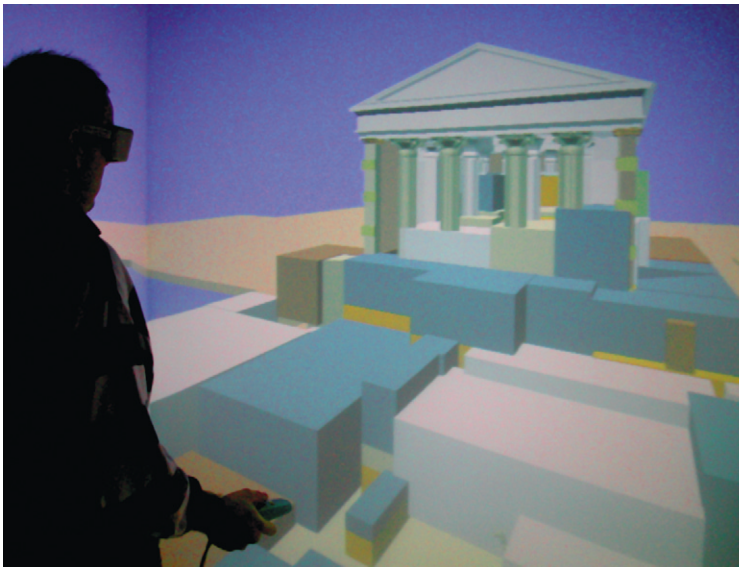

Fig. 4. Archave image from the full reconstruction version of archave.

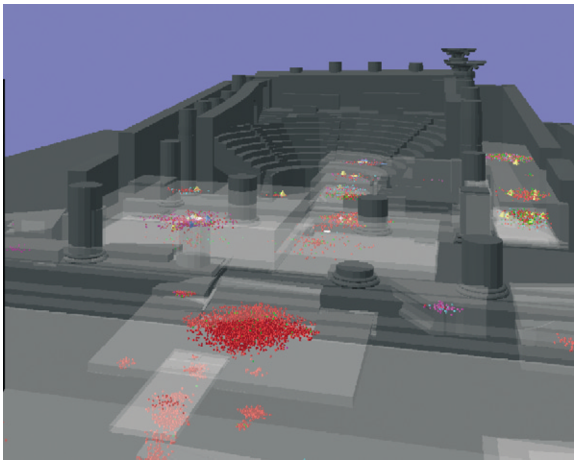

Fig. 5. Archave representation of excavation site with artifacts.

Scale and navigation: Another research issue concerns the intertwined matters of scale and navigation. Should we work at full scale? In miniature? What are the tradeoffs for archaeologists studying a site like Petra that is the size of three football fields? For a scale model that fits within the Cave, navigation can be primarily via body motions. Larger scales require longer-distance navigation, and the virtual world must move relative to the Cave, obviating any fixed mental model a user may have created. Motivated in part by these multiscale navigation needs of Archave, Joe LaViola developed step-WIM navigation [58], in which the user employs gestures to access a world-in-miniature for navigating large distances and familiar body-motion navigation for shorter distances.

Productivity: Problems arising from the ephemeral nature of the Cave are an active area of research [59–61]. Most desktop productivity applications create persistent artifacts—word processors and editors produce documents, modelers create geometric models, and spreadsheets create analysis documents. The ephemeral nature of Cave experiences limits their scientific utility. Archave, for example, has been primarily a browser of the archaeological record, and while archaeologists have found this valuable, they need to capture their work in the Cave for later analysis. We are exploring several approaches to this issue, some of which are described below in Section 5.4.

Quantitative evaluation: Quantifying the value of IVR for this application is difficult. Initially, we collected anecdotal evidence during demos to archaeologists. We then attempted to design a user study, defining quintessential tasks as quantitative performance yardsticks. We failed. Such user studies require a quick task because it must be repeated many times. But a task must also be representative of what real users will do. In our scientific applications, the real tasks are still being identified and so are not clearly defined. Those that have been identified tend to be complex and time-consuming, e.g., determining relative dating of a set of spatially related finds from several trenches. Defining tasks for quantitative assessment is thus difficult because of all the competing constraints. For now we compromise and give structured 1–2 hour "demos" that we record on video and analyze anecdotally [52]. As this research issue extends to all of our applications, our continuing research agenda includes a search for more effective ways to evaluate user interface concepts in IVR.

Discipline-specific VR sensitivity: Another second limiting factor in Cave utility was the sensitivity of Archave users, seemingly more than those in some other disciplines, to fatigue and discomfort in the Cave. This sensitivity seems to be shared by many artists as well. This result is surprising, running counter to an expectation of little disorientation because of the "familiar" and simple spatial nature of the depiction, and must be further investigated.

5.3. Bioflow visualization

Our second scientific application, in collaboration with bioengineer Peter Richardson of the Division of Engineering, CFD specialists George Karniadakis and Igor Pivkin of Applied Math, and IVR specialists Andrew Forsberg, Bob Zeleznik, and Jason Sobel of Computer Science, is a virtual environment for visually exploring simulated pulsatile blood flow within coronary arteries [62]. Cardiologists and bioengineers hypothesize that characteristics of the flow within these vessels contribute to the formation of plaques and lesions that damage the vessels. A significant fraction of the population suffers from this problem. We conjecture that IVR is an appropriate way to develop a better understanding of the complex 3D structure of these flows.

In Fig. 6, a scientist uses our system to study simulated pulsatile flow through a model of a bifurcating coronary artery; lesions typically form just downstream of these splits. IVR offers a natural exploration environment. The representation in Fig. 7 shows an overview of all of the data available as a starting point for exploration. This "synoptic" view is analogous to synoptic weather diagrams, which simultaneously show geography, temperature, pressure, precipitation, etc. In our representation the wire-mesh artery walls show quantities on them via variably distributed splats, yellow for pressure and green for residence time, and the flow is represented within the artery with particle paths that advect through the flow.

To make the view both fast enough to compute and possible to interpret, particle paths are concentrated in "interesting" areas. Only a subset of the possible paths is visible at any one time, permitting a faster frame rate. By cycling through which particle paths are visible, however, all the time-varying flow can be shown over a period of viewing time. The application also supports some interaction with the flow, including placement of persistent streaklines, uniform coloring or deletion of all particles that pass through a specified region, and controls for the rate, density, and for defining the "interesting" areas to emphasize.

The synoptic view provides a starting point for exploration that gives more insight than an empty space. Other parts of the interface give a way to explore the synoptic view, adjust it, and annotate features as they are discovered. Thus far, we have been able to understand the simulation process that creates our flow data better than has been possible with traditional workstation-based tools such as Tecplot [63]. We are just starting to find new arterial flow features.

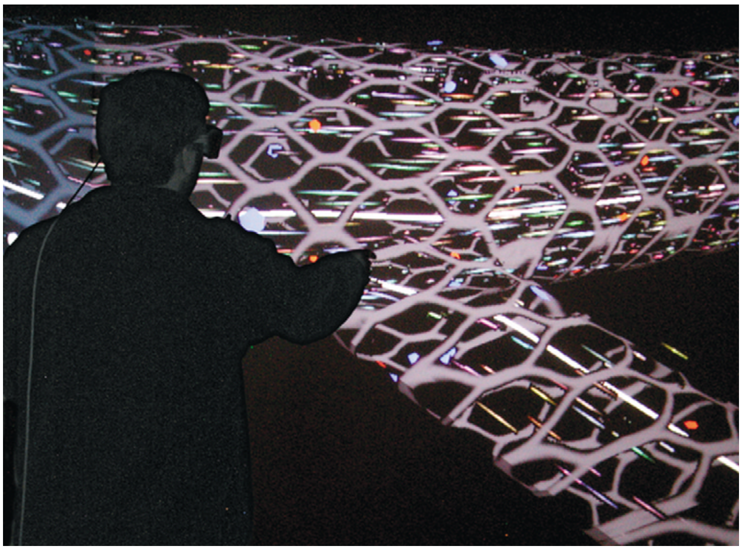

Fig. 6. Bioengineer studying pulsatile flow within an idealized virtual model of a bifurcating coronary artery. Viewing this scene with head-tracked stereo glasses causes particles and the textured tubular vessel wall to "jump" into the third dimension and be much more clear.

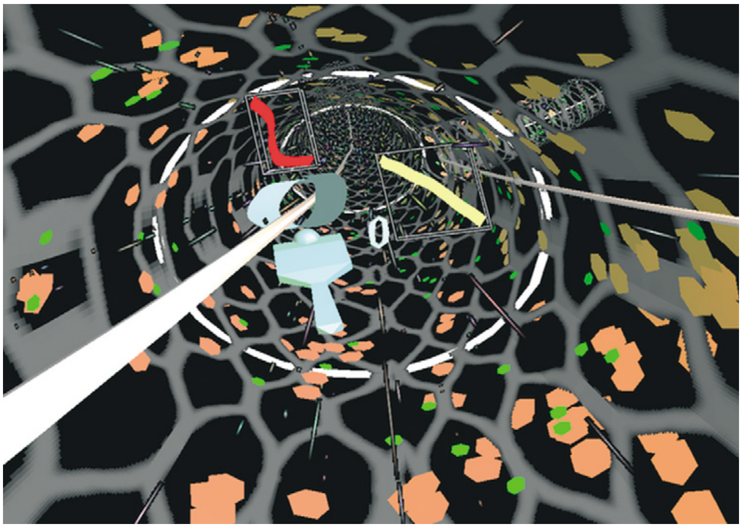

Fig. 7. Synoptic view of interior of the artery, including streaklines, annotations, sponges, pierce planes, pressure splats, and particle paths.

5.3.1. Research issues

Beyond the perennial issues of frame rate, tracker lag and calibration, and model complexity, as discussed in [2], some more subtle ones have emerged as we developed this application.

Multivalued visualization: As in the Archave project, this application has more different types of data than we can visualize at once. In addition to the velocity within the artery, which is a 3D-vector field, pressure and residence time are defined on the walls of the vessel. Beyond those quantities, our fluid flow collaborators want to look at pressure gradient, vorticity, other derived quantities in the flow, and the structure of critical points.

Our synoptic visual approach for displaying this multivalued data is partly motivated by Interrante's work [64] demonstrating that patterns on nested surfaces are more effective at conveying shape than transparency. We carry that to a volumetric display of particles and also use motion to increase the apparent "transparency". We also build upon the layering of strokes introduced in [33,56,65]. With these principles we can display all the data with very little occlusion as a starting point for study of the flow. We are also exploring other more feature-based visual abstractions for the flow showing, for example, critical points and their relationships, coherent flow structures, or regions of the flow that may be separating. They have the potential to abstract the essence of a flow more efficiently, although at the risk of missing important but unexpected flow behavior.

What makes a visual representation effective? For understanding steady 2D flows, important tasks that are simple enough for a user study include advecting particles, locating critical points, and identifying their types and the relationships among them [57]. The same tasks are likely to be relevant in 3D but there are likely to be other simple tasks that we have not identified yet, and there are certainly complex tasks that are beyond the scale appropriate for quantitative statistical evaluation needed for a usability study. The discovery of such measures of efficacy will clearly influence the design of visualization methods.

Scale and navigation: While the scale and navigation issues are somewhat similar to those of Archave, they differ in that Archave has a natural life-size scale while arterial blood flow requires a more "fantastic voyage" into the miniature. When should our artery model be life-size? Large enough to stand in? Somewhere in between? Under user control? Any one such specific question could probably be answered given some context and a user study. But would the results generalize to different navigation strategies or to a different visual representation? For our synoptic visualization, a model about 6 ft in diameter seems to give the best view of flow structures of interest from inside. Smaller-scale models are too difficult to get one's eyes inside. We have explored several navigation metaphors, including wand-directed flying, direct coupling of the virtual model to the hand, and a railroad-track-like metaphor with a lever for controlling position along a path down the center of the artery. Each has strengths and weaknesses; none is ideal in all situations. Design issues like these of scale and navigation, both at a specific and at an integrative level, have become an important part of our ongoing research.

Our group at Brown is working with half a dozen faculty from the Rhode Island School of Design (RISD) to address some of these issues, a process reminiscent of Cox's "Renaissance teams" [66]. Our goals are threefold: first, to develop new visual and interaction methods; second, to explore the design process; and third, to develop a curriculum to continue this exploratory process. For the last four months we have held a series of design sessions and have built a common frame of reference among the designers, computer graphics developers, and domain scientists; we are ready to proceed with new designs. We will offer a Brown/RISD course in Fall 2002 to 6–10 students from each institution.

Design iteration costs: One of the highest costs of developing IVR applications is iterating on their design. Archave took almost three years going through four significant design phases. For our bioflow application, we have implemented three significantly different designs in about 18 months. Parts of this process are likely to be essential and incompressible, but we believe that other parts are accidental, to use Brooks' terminology [67]. A Brown graduate student, Daniel Keefe, is working together with our RISD collaborators to find ways of speeding it up. Initial efforts build on his CavePainting application [68], which is being used to help quickly sketch visual ideas in the Cave.

Productivity: As with Archave, scientists would be more productive if they had tools for annotating their discovery process, facilities for taking results away, and some way to build on earlier discoveries. We are exploring a number of possible alternatives that would enable researchers to take their notes and versions of the applications away from the Cave environment.

5.4. Brain white-matter visualization

In a collaborative biomedical effort, developed by computer science graduate students Cagatay Demiralp and Song Zhang in collaboration with, among others, neurosurgeons from MGH and brain researchers from NIH we are developing a visualization environment for exploring a relatively new type of medical imaging data: diffusion tensor images (DTI). Humans are 70% water. The structure of our tissues, particularly in the nervous system, influences how that water diffuses through the tissues. In particular, in fibrous structures like muscle and nerve bundles, water diffuses faster along the fibers than across them. DTI can measure this directional dependence. Measurements of the diffusion rate have the potential to help understand this structure and, from it, connectivity within the nervous system. This will help doctors better understand the progression of diseases and treatment in the neural system. DTI provides volume images that measure this rate of diffusion at each point within a volume. Viewing these volumes is a challenge because the measurement at each point in the volume is a second-order tensor consisting of a 3×3 symmetric matrix containing six inter-related scalar values.

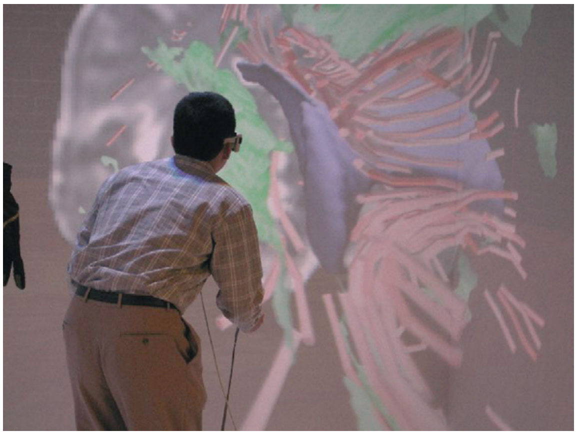

Fig. 8. Brain model use in Cave.

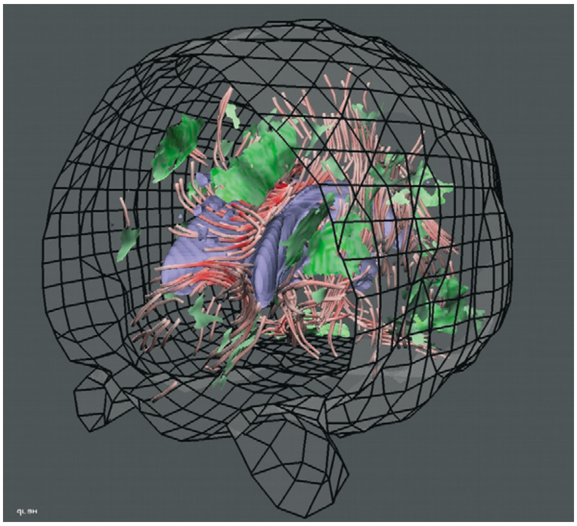

Fig. 9. Brain model in wireframe enclosure.

We generate geometric models with the volumes to represent structures visually and display them in the Cave. In Figs. 8 and 9, an abstraction derived from a DTI of a human brain, different kinds of geometric models represent different kinds of anatomical structures; the red and white tubes represent fibrous structures, like axon tracts; the green surfaces represent layered structures, like membrane sheets or interwoven fibers; the blue surfaces show the anatomy of the ventricles.

Our experience thus far is that IVR facilitates a faster and more complete understanding of these complicated models, the medical imaging data from which they are created, and the underlying biology and pathology. Applications we are pursuing to explore this claim include preoperative planning of tumor surgery, quantitative evaluation of tumor progression under several conditions, and the study of changes due to surgery on patients with obsessive compulsive disorder.

5.4.1. Research issues

Once again, frame rate, lag, visual design, and interaction design are intertwined issues. Biomedical researchers using this application almost always want more detail and will tolerate frame rates as low as 1 FPS for visualization, even knowing that they can move their heads only very slowly so as to avoid cybersickness. This clearly detracts significantly from the feelings of immersion and presence. Worse, many interactions are virtually impossible at that rate, so we struggle to balance these conflicting requirements.

Multivalued visualization: These volume-filling second-order tensor-valued medical images consist of six values at each point of a 3D volume. Often, we have additional co-registered scalar-valued volumes. We continue to search for better visualization abstractions through close collaboration with scientists, since it is only thus that we can define what is important and what can be abstracted away. Exploring differences between subjects or longitudinally within one subject brings the additional challenge of simultaneously displaying two of these datasets [57,69].

Design: Beyond the data visualization design, we notice again, as with Archave, the importance of virtual context. Working within a virtual room, created by providing wall images within which the brain model is placed, gives users a subjectively more compelling experience. Several users report that the stereo seems more effective and that they can "see" better with the virtual walls around them. In collaboration with perceptual psychologists, we are exploring why.

Scale and navigation: As with Archave, full-scale seems a natural choice. However, much as with our arterial flow visualization users report a preference for larger scales. At full scale, users end up struggling to see small features. They can make the projected image larger by moving closer to the virtual object, but then have to squint and strain; also, they find it difficult to fuse the stereo imagery when features are up close. With the approximately 2:1 scale models used thus far, exploration has been almost exclusively from the outside looking in, unlike the flow visualization, where most of the exploration has been from within the flow volume. Could each area benefit from interaction techniques that the other is using? Are there intrinsic differences?

Productivity: We continue to see productivity issues within the Cave. It is often run in batch mode, with time slots, limited access, and a (probably ineffective) urgency to be "efficient". Users want a personal Cave (PC) to sit inside and think, with the computer doing nothing. Some of our most effective sessions have involved users working for a while, then sitting on the floor of the Cave talking, then stepping outside to look at things, then going back in to look some more.

Collaborating in a single Cave has proven very important—we almost always have at least two domain experts in the Cave at a time so that we can listen to them talk to each other. One disadvantage is that they cannot point with real fingers because only one viewer is head-tracked. In some ways, long-distance collaboration solves this problem because each viewer has their own Cave, but that introduces other problems of synchronization and communication. Users of this application were the first to ask for physical objects to augment the virtual environment—anatomy books and printed medical images, for example. These are natural and familiar objects for neuroanatomists to work with and can be more easily used when not virtual. Our current practice is to iterate moving out of the Cave to work with reference books, keeping the image on the front wall of the Cave for reference, then returning to the Cave for further exploration.

5.5. Mars terrain Exploration

In the fourth and final IVR application discussed here, terrain from Mars is modeled for planetary geologists so they can "return to the field". The predominant means of visualizing satellite data is 2D imaging (at multiple resolutions) with color maps representing elevation and other terrain attributes.

However, Mars Orbiter Laser Altimeter (MOLA) missions give a third dimension to the terrain. Incorporating topography information into analysis opens new avenues for creating new, modifying, or ruling out existing hypotheses about geological processes. Our primary goal is to create a complete model integrating all of the available data that is interactively viewable. The geologists' ultimate goal is to understand the geological processes on Mars from the current surface structures, including the relationships among the strata that are partially visible on the surface. More pragmatically, studying the terrain gives an opportunity to search for sites that warrant further study and to identify potential landing sites for future missions.

In Fig. 10, James Head of Brown's GeoSciences Department uses a handheld IPAQ PDA, which provides convenient interactive input to control global position and rendering styles while flying over the terrain of Mars. Mars data available to the exploration process include elevations from the MOLA and color images of surface swathes from the Mars orbiter camera (MOC). Altitude information we are displaying currently is on a 7200×7200 grid. The thousands of color swathes can be as large as 6000×2000 pixels and cover a small region of the surface.

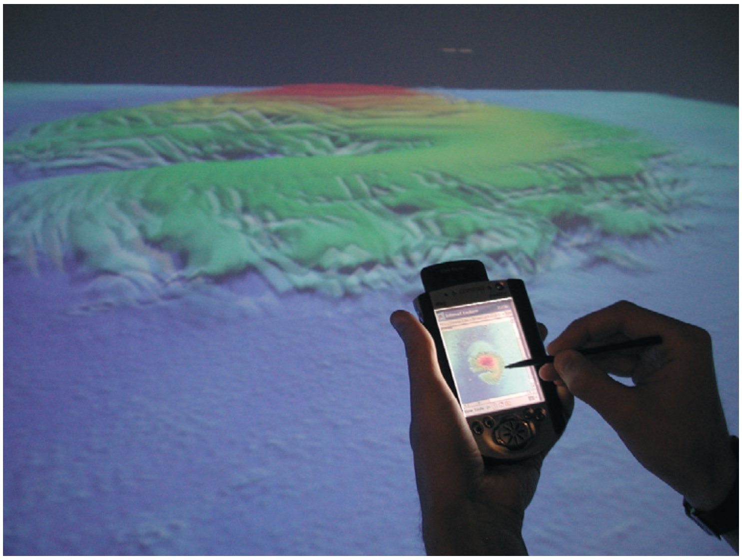

Fig. 10. Geologist using PDA to aid Mars flyover.

5.5.1. Research issues

Scale and navigation: What is the best scale for Mars study? At full scale, navigation becomes extra-planetary and unfamiliar. Smaller scales appear to be more appropriate, although the scale that we currently use requires hybrid navigation: flying for most movement coupled with a much smaller-scale map for taking larger steps. As with many of the other applications, users are more comfortable driving than being driven, particularly moving backwards.

Performance: One of the most significant limitations for this application is performance. A consistent 30 + frames per second is very difficult to maintain. A naïve approach to rendering the 7200×7200 heightfield yields about 100 million triangles!). Lawrence Livermore National Laboratory's (LLNL) real-time optimally adapting meshes (ROAM) [70] software for view-dependent terrain rendering improves performance, but more progress is necessary. The geologists consistently request more detail, exacerbating the problem.

6. Tele-immersion: an animating vision

As we observed in the Introduction, computing increasingly emphasizes collaboration, although the tools for effective collaboration beyond simple shared whiteboards are still immature, even for conventional desktop environments. While IVR itself is also still an immature technology, there is great societal need to combine these two immature technologies for tele-immersion. Furthermore, tele-immersion provides the R&D community with a very demanding "driving application" for all the enabling technologies involved.

Leigh et al. [44]define collaborative VR as having its goal to "reproduce a face-to-face meeting in minute detail". They then say that "tele-immersion moves beyond this idea, integrating collaborative VR with audio- and video-conferencing that may involve data mining and heavy computation". Our definition of tele-immersion is similar, but we augment the notion of traditional audio- and video-conferencing, possibly with synthetic avatars, with real-time reconstruction of remote scenes and their occupants for a greater presence. As Jaron Lanier states [43] "It combines the display and interaction techniques of VR with new vision technologies that transcend the traditional limitations of a camera. Rather than merely observing people and their immediate environment from one vantage point, tele-immersion stations convey them as 'moving sculptures' without favoring a single point of view." The effect is as much as possible like looking through glass walls at other environments that are blended with one's own.

6.1. UNC's office of the future

This vision is illustrated in Fig. 11, which shows Henry Fuch's office as he engages in the collaborative design of a new head-mounted display. Noteworthy features include the "sea of cameras" and projectors in the ceiling. Several prototype versions of this environment have been built and demonstrated under NSF and Advanced Networks & Services sponsorship, as described in [43], through collaborations among UNC, University of Pennsylvania, Brown, and Advanced Networks & Services. In these prototypes the projectors are used to create the real-time stereo display; in the future they may help scene acquisition by displaying

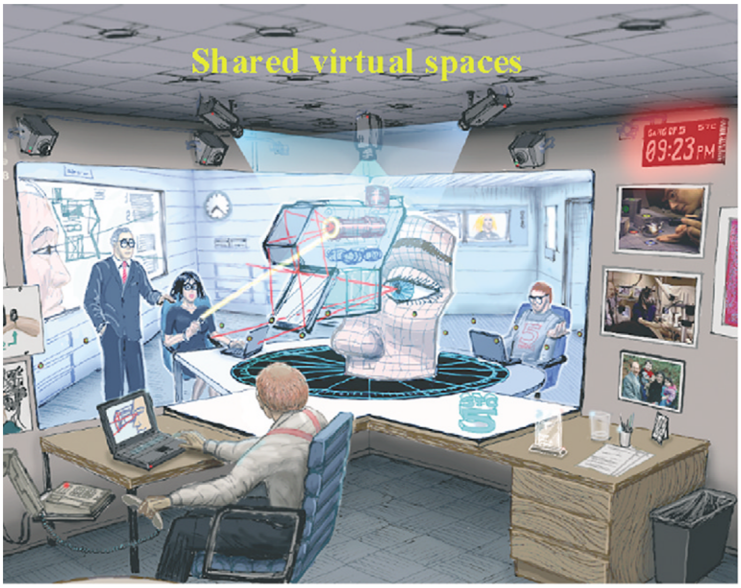

Fig. 11. UNC's office of the future vision. Image courtesy of Andre State of UNC-Chapel Hill.

imperceptible structured light patterns [71] that help the reconstruction algorithms create disparity and depth maps from the camera information. The display is created by combining computer-generated imagery with reconstructions from remote sites (shown as very dense confetti clouds, subject to inevitable errors, dropouts, and noise). One could also blend in high-resolution data from image-based rendering algorithms.

To date, using only seven cameras and commodity PC hardware, the proof-of-concept demonstrations have shown small-volume reconstructions with two or three sites at several frames per second. A collaboration among UNC, UPenn, and the Pittsburgh Supercomputing Center aims to improve both the reconstruction volume and the spatio-temporal resolution through using many more cameras and massive parallel computation. While the imagery in these early experiments is crude, the sense of presence is very compelling because of the real-time nature of the display and the ability to observe the remote scene as a genuine 3D scene using head tracking. The collaborative interaction with synthetic models in the shared space also aids the illusion. In another experiment, (Fig. 12), UNC laser-scanned a static room with a mannequin over 20 min. The batch reconstruction was so gratifyingly realistic that a demoee in head-tracked stereo glasses said it felt as if "a chain saw had been used to cut a hole in the wall".

Fig. 12. Tele-immersion experiment with UNC and the University of Pennsylvania.

6.2. ETH-Zurich blue-c project

Marcus Gross and colleagues in CS and Engineering at ETH-Zurich [72] are engaged in another ambitious project to create shared virtual environments with realistic reconstructions of humans, the blue-c project. This project simultaneously combines the acquisition of live video streams with the projection of virtual reality scenes. Color representations with depth information about the users will be generated from the multiple video streams in real time.

A blue-c portal currently consists of three rectangular projection screens, much as in a Cave. The novelty of the blue-c approach is that these are glass screens containing a liquid crystal film. These screens can be switched from a whitish translucent state (for projection) to a transparent state (for acquisition by cameras located outside the portal facing the screens.). The projection technology is active stereo based on two LCD projectors per screen, which are synchronously shuttered with the stereo glasses and the acquisition hardware; the reconstruction uses the image-based visual hull algorithm [73]. The modification of shutter glasses to accommodate both projection and picture acquisition is intended to handle the opposing requirements of darkness for image projection and light for image acquisition, thereby enabling the use of inexpensive camera acquisition systems in a tele-immersive collaboration environment.

6.3. *Electronic books for teaching surgical procedures*

Following on the UNC–UPenn–Brown tele-immersion work described above, the NSF has funded a three year grant to Brown and UNC for immersive electronic books that in effect blend a "time machine" with 3D hypermedia. This will add an additional important dimension to tele-immersion, the ability to record experiences: a viewer will thus be immersed in the 3D reconstruction [74] and will be able literally to walk through the scene or move backward and forward in time. Our research will focus the driving application of teaching surgical management of difficult, potentially lethal injuries.

Our aim is to develop a new paradigm for teaching surgical procedures that allows surgeons to witness and explore (in time and space) a past surgical procedure as if they were there, with the added benefit of instruction from the original surgeon or another instructor, as well as integrated 3D illustrations, annotations, and relevant medical metadata. The trainees should be able freely and naturally to walk around a life-sized, high-fidelity, 3D graphical reconstruction of the original time-varying events, pausing or stepping forward and backward in time to satisfy curiosity or allay confusion. To make this demanding vision a reality, we are bringing together experts in several disciplines to leverage their prior work in tele-immersion, time-varying 3D scene capture, interaction metaphors, "cinematic" techniques and authoring tools (Fig. 13).

Since the full extraction functionality is not yet available, in the short term we are working with a hand-made model of a surgery that includes off-the-shelf 3D anatomy objects. As a bridge between our short- and long-term work, we are creating synthetic video streams from which 3D geometry will be extracted. In this case, we will create the synthetic video streams directly from the hand-made model.

6.3.1. *Research issues*

The research issues arising out of current work in tele-immersion encompass a wide range of technical and social issues, ranging from the ever-present needs for latency reduction and compensation, faster frame rates, more reliable and accurate tracking, to information and user interface design, and social issues such as privacy. A successful outcome requires cross-disciplinary work with computer vision researchers as well as collaboration among a variety of hardware and software disciplines.

7. Conclusion

While IVR and scientific visualization have their own significant histories and accomplishments as separate fields, their intersection is less common and considerably more immature. Because scientific visualization so often involves handling large datasets, it is a great stressor of all aspects of IVR technology. To create satisfactory user experiences, IVR's requirements of low latency and

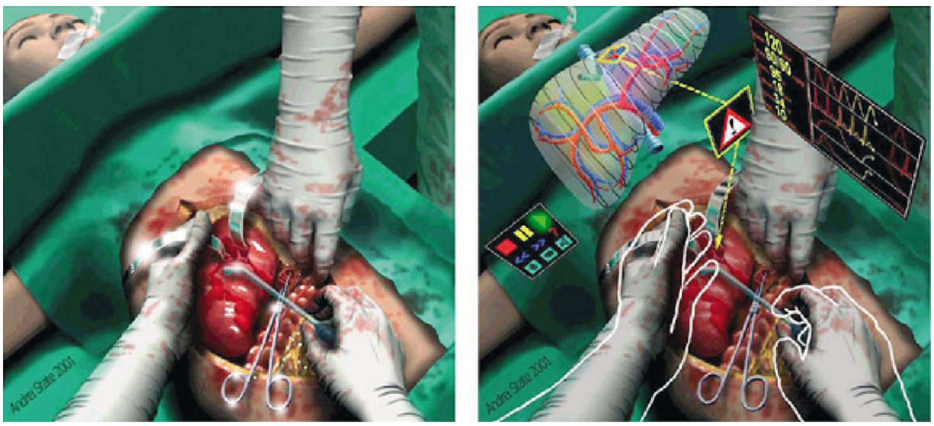

Fig. 13. Electronic surgery textbook.

high frame rate, and high-resolution rendering and interaction handling (e.g., head and hand tracking) are much more stringent than those for workstation graphics. This may force severe restrictions in the complexity of models and datasets that IVR can handle adequately. Such constraints put a serious burden of proof on IVR proponents to show that the advantages of the immersive experience outweigh the disadvantages of higher cost, of scheduling a one-of-a-kind institutional facility such as a Cave, and, most important, of complexity restrictions in what can be visualized.

While there is production use of IVR for scientific visualization, e.g., for drug design, oil exploration, and nuclear weapon simulations, most of it is still experimental. Indeed, for many working scientists and engineers, projection displays and high-resolution tiled displays such as PowerWalls may provide adequate, indeed preferable solutions for their current visualization problems. Nonetheless, we remain hopeful that with expected price/performance improvements in rendering technology and interaction devices, as well as in software, there will be visualization applications for which the additional insight that the immersive experience provides will make IVR the most satisfactory mechanism. We are especially interested in the promise of the most "extreme" and most difficult form of IVR, collaborative tele-immersion.

There is a spectrum in the degree of certainty with which we may anticipate continued improvement in various aspects of IVR technology. Most certain is the faster-than-Moore's-law improvement in rendering engines, driven largely by the commoditization of PC graphics cards and consoles for the game industry. Scalable graphics clusters of PCs with 3D graphics cards are now becoming standard, replacing the much more expensive and specialized high-end workstation-based renderers [75]. There is less steady progress in new types of display surfaces (e.g., comfortable, high-resolution, large-field-of-view headmount displays), although personal IVR devices such as projection-based semi-immersive workbenches and personal desk-sized hemisphere displays [76] are advancing steadily.

Especially troublesome is the very slow progress towards more comfortable (e.g., very lightweight, small, and untethered) and more robust interaction devices with higher spatio-temporal resolution. Minimizing or compensating for tracker latency is an important problem for both local tracking and in collaborative applications. Prediction algorithms provide a valuable tool for addressing this problem. At Brown, Joe LaViola is working on a framework for assessing the value of various predictive tracking algorithms for tasks such as tracking head and hand motion. Other basic interaction technologies such as speaker-independent continuous speech recognition, particularly in noisy environments, still need work, not to mention the very difficult problem of providing distributed haptics, especially haptics for mobile users.

Even less far along and requiring even more fundamental research are several areas key for handling very large models and datasets. Among these are:

(1) Culling and summarizing algorithms (embodied in semi-autonomous agents) to let users focus on the small percentage of data that is typically worth examining in detail. Traditional scientists and engineers must work with knowledge engineers and other experts in AI and searching/data mining techniques to encode pattern-recognition processes that detect characteristic features, anomalies and errors in their typical datasets. Ultimately, as part of the human/computer partnership, an agent would cull through the data and point out features of potential interest to the scientist, much as a colleague would. As shown repeatedly in designing expert systems, it is hard both to learn how experts perform their tasks and to design good rules and algorithms that mimic their behavior.

(2) Visualization techniques for higher-dimensional (time-varying) data, including the interactive exploration, annotation and recording, and immersive playback of those visualizations. As mentioned above, we expect significant inspiration and help from traditional art and design disciplines.

(3) A new and highly interdisciplinary design discipline for creating compelling immersive environments, including visualizations. We cannot expect interdisciplinary design teams to evolve spontaneously or to invent their techniques from scratch each time. And environment design certainly cannot be left to traditionally-educated computer scientists, most of whom have no formal background in the creative arts and design or in human perceptual, cognitive and social systems. It is encouraging that several schools are taking seriously the challenge of educating designers of interactive and immersive experiences. We can see the beginnings in programs in entertainment science, typically focused on game design, related to but certainly not the same as the design of virtual environments [77–80]. Even more than UI design, immersive environment design will be a highly interdisciplinary design discipline that rests on foundations in the human sciences, computer science and engineering, traditional design disciplines such as graphic and industrial design, story-telling disciplines such as theater and movies, and interactive experiences, especially of the massive multi-player variety.

In conclusion, we fall back on the cliché that this field is largely still in the Kitty-Hawk stage of preliminary experimentation, with the equivalents of such benefits of modern airplane technology as ubiquitous airtravel, airmail, and air freight still mostly in the future. What is especially lacking today is a virtuous commoditization cycle that makes the hardware and software necessary for IVR robust and affordable on a wide scale. While there are daunting problems to be overcome, our experience and that of other experimenters lead us to believe that the experience of immersion is sufficiently more powerful and more tuned to the human sensorium than is conventional interaction with computers that investments in the field must continue, indeed accelerate.

Acknowledgements

We happily acknowledge the contributions of faculty, students, and staff who worked on the Brown University research projects described herein, as well as Fred Brooks and Andy Forsberg who gave us substantive input and/or feedback on the article to help improve it. As always we are deeply grateful for the superb editing of Katrina Avery. We also are grateful for the support of our sponsors, including NSF, through the NSF Science and Technology Centre for Computer Graphics and Scientific Visualization (whose members are Brown, Caltech, Cornell, UNC Chapel Hill and the University of Utah), the Department of Energy's ASCI program, IBM Corporation, Microsoft Research, and Sun Microsystems.

References

[1] Carolina Cruz-Neira, Daniel J Sandin, Thomas A DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In: Proceedings of the ACM SIGGRAPH '93, 1993.

[2] Andries van Dam, Andrew S Forsberg, David H Laidlaw, Joseph J LaViola, Jr., Rosemary M Simpson. Immersive VR for scientific visualization: a progress report. IEEE Computer Graphics and Applications 2000;20(6):26–52.

[3] Andries van Dam. User interfaces: disappearing, dissolving, and evolving. Communications of the ACM 2001;44(3):50–52.

[4] Raj Reddy. Will 3G supercomputers also gather dust? In: Richard F. Rashid (ed.), CMU computer science: a 25th anniversary commemorative, ACM Press, 1991 ISBN 0-201-52899-1.

[5] Raj Reddy. The challenge of artificial intelligence. IEEE Computer 1996;29(10): 86–88.

[6] Digital Light Processing, a Texas Instruments company: http://www/dlp.com/

[7] Cambridge Display Technology, Ltd., LEP developer http://www.cdtltd.co.uk/. For an overview of organic electroluminescence and light-emitting polymers, see: Kaminsky, Kurtis. Organic Electroluminescence and Light-Emitting Polymers. www.ualberta.ca/~kaminsky/organicelectroluminescence.htm.

[8] Hector Garcia-Molina, Yannis Papakonstantinou, Dallan Quass, Anand Rajaraman, Yehoshua Sagiv, Jeffrey Ullman, Jennifer Widom. The TSIMMIS project: integration of heterogeneous information sources. In: Journal of Intelligent Information Systems 1997.

[9] Microsoft. NET: http://www.microsoft.com/net/defined/whatis.asp

[10] IBM WebSphere: http://www-3.ibm.com/software/info1/websphere/index.jsp

[11] Sun Microsystems Sun ONE: http://www.sun.com/software/sunone/

[12] Tobias Hollerer, Steven Feiner, Drexel Hallaway, Blain Bell, Marco Lanzagorta, Dennis Brown, Simon Julier, Yohan Baillot, Lawerence Rosenblum. User interface management techniques for collaborative mobile augmented reality. Computers and Graphics 2001;25(5): 799–810.

[13] Jonathan Grudin. Computer-supported cooperative work: history and focus. IEEE Computer 1994;27(5):19–26, http://www1.ics.uci.edu/~grudin/Papers/IEEE94/IEEEComplastsub.html.

[14] University of Minnesota PowerWall project: http://www.1cse.umn.edu/research/powerwall/powerwall.html

[15] Elizabeth F Churchill, David N Snowdon, Alan J Munro editors. Collaborative virtual environments: digital places and spaces for interaction. New York: Springer, 2001. ISBN: 1-85233-244-1.

[16] Advanced Network and Services, Inc. NTII (National Tele-Immersion Initiative) project: http://www.advanced.org/tele-immersion/

[17] William Buxton. Less is more (more or less): some thoughts on the design of computers and the future. In: Denning P, editor. The invisible future: the seamless integration of technology in everyday life. New York: McGraw-Hill, 2001.

[18] Philip R Kennedy, Roy A E Bakay, Melody M Moore, Kim Adams, John Goldwaithe. Direct control of a computer from the human central nervous system. IEEE Transactions on Rehabilitation Engineering 2000; 8(2):198–202.

[19] Mark Beaulieu. Wireless internet applications and architecture: building professional wireless applications worldwide. Reading, MA: Addison-Wesley, 2002 ISBN 0-201-73354-4.

[20] Oliver Bimber, L Miguel Encarnacao, Andre Stork. A multi-layered architecture for sketch-based interaction within virtual environments. Computers and Graphics 2000;24(6): 851–867.

[21] Philip R Cohen, Michael Johnston, David McGee, Sharon Oviatt, Jay Pittman, Ira A Smith, Josh Clow. QuickSet: multimodal interaction for distributed applications. In: Proceedings of the ACM Multimedia '97, 1997. p. 31–40.

[22] Stanford University HCI and Engelbart history site: http://sloan.stanford.edu/MouseSite/1968Demo.html

[23] NetMeeting home page: http://www.microsoft.com/windows/netmeeting/default.asp

[24] Squeak open source language, Nebraska application: http://minnow.cc.gatech.edu/squeak/1356

[25] Norbert A Streitz, Jorg Geisler, Torsten Holmer, Shin'ichi Konomi, Christian Muller-Tomfelde, Wolfgang Reischl, Petra Rexroth, Peter Seitz, Ralf Steinmetz. i-LAND: an interactive landscape for creativity and innovation. In: Proceedings of the ACM SigCHI '99, 1999. p. 120–127.

[26] William Buxton, George Fitzmaurice, Ravin Balakrishnan, Gordon Durtenbach. Large displays in automotive design. In: IEEE Computer Graphics and Applications 2000; 68–75, July/August 2000. http://www.billbuxton.com/largeDisplays.pdf
[27] Roger J Hubbold. Collaborative stretcher carrying: a case study. In: Proceedings of the ACM Eurographics Workshop on Virtual Environments, 2002. p. 30–31.
[28] Adnan Anser, Denilson Rodrigues, Jaydev P Desai, Kostas Daniilidis, Vijay Kumar, Mario F M Campos. Visual and haptic collaborative tele-presence. Computers and Graphics 2001;25(5):789–798.
[29] Ray Bradbury. The veldt [“The World the Children Made”]. In :The Saturday Evening Post, September 23, 1950.
[30] Ivan Sutherland. The ultimate display. In: Information processing 1965: Proceedings of IFIP Congress 65, vol. 2, 1965. p. 506–508.
[31] Ivan Sutherland. A head-mounted three dimensional display. In: Proceedings of the 1968 Fall Joint Computer Conference. AFIPS Conference Proceedings, Part I, vol. 33, 1968. p. 757–764.
[32] Michael Meehan, Brent Insko, Mary Whitton, Frederick P Brooks, Jr. Physiological measures of presence in virtual environments. In: Proceedings of the 4th Annual International Presence Workshop, Philadelphia, PA, 2001. http://nimbus.ocis.temple.edu/~mlombard/P2001/papers.htm
[33] R Michael Kirby, Haralambos Marmanis, David H Laidlaw. Visualizing multivalued data from 2D incompressible flows using concepts from painting. In: Proceedings of the IEEE Visualization ’99, October 1999. p. 333–340.
[34] Randy Pausch, Jon Snoddy, Robert Taylor, Scott Watson, Eric Haseltine. Disney’s aladdin: first steps toward storytelling in virtual reality. In: Proceedings of the ACM SIGGRAPH ’96, August 1996.
[35] John Deere VR. Research with Iowa State: http://www.vrac.iastate.edu/research/prototyping/deere/
[36] General Motors vehicle design: http://www.gm.com/company/news_events/events/concept_vehicles/virback.htm
[37] SpiderWorld for phobias: http://www.hitl.washington.edu/research/exposure/
[38] SnowWorld: http://www.wired.com/news/print/0,1294,42084,00.html
[39] Spider World for pain management: http://www/hitl.washington.edu/research/burn/
[40] Virtual Reality Surgery Report: http://www.msnbc.com/news/281055.asp?cp1=1
[41] Steven Haker, Sigurd Angenent, Allen Tannenbaum, Ron Kikinis. Non-distorting flattening for virtual colonoscopy. In: Proceedings of the MICCAI (Medical Image Computing and Computer-Assisted Intervention), 2000. p. 358–366.
[42] Desa Philadelphia. Video traveler. Time 2002;159(6): Y2–Y4.
[43] Jaron Lanier. Virtually there. Scientific American 2001; 284(4):2–11.
[44] Jason Leigh, Andrew E Johnson, Maxine Brown, Daniel J Sandin, Thomas A DeFanti. Visualization in teleimmersive environments. IEEE Computer, 1999;66–73, http://www.startap.net/images/PDF/Leigh_IEEEComputer.pdf
[45] EVL’s CAVERN: http://www.evl.uic.edu/cavern/
[46] EVL connection of two conferences via CAVERNsoft and STAR TAP: http://www.ncsa.uiuc.edu/News/Access/Releases/00Releases/001115.STARTAP.html June 7, 2002.
[47] VisualEyes project: http://evlweb.eecs.uic.edu/research/template_res_project.php3?indi=134
[48] Fred P Brooks, Jr. The computer scientist as toolsmith II. Keynote/newell award address at SIGGRAPH 94, Orlando, July 1994, in Communications of the ACM (CACM), 1994;39(3):61–68.
[49] Fred P Brooks, Jr. The computer ‘Scientist’ as toolsmith: studies in interactive computer graphics. In: Proceedings of the IFIP (International Federation of Information Processing) Congress ’77, August 1977.
[50] Deborah Hix, J Edward Swan II, Joseph L Gabbard, Mike McGee, Jim Durbin, Tony King. User centered design and evaluation of a real-time battlefield visualization virtual environment. In: Proceedings of the IEEE Virtual Reality ’99, 1999. p. 96–103.
[51] Brown University’s Center for Digital Shape—Archaeology Photogrammetry—Entropy (S.H.A.P.E.): http://www.lems.brown.edu/vision/extra/SHAPE/
[52] Daniel Acevedo, Eileen Vote, David H Laidlaw, Martha S Joukowsky. Archaeological data visualization in VR: analysis of lamp finds at the great temple of Petra, a case study. In: Proceedings of the IEEE Visualization 2001, October 2001.
[53] David B Cooper, Andrew Willis, Stuart Andrews, Jill Baker, Yan Cao, Dongjin Han, Kongbin Kang, Weixin Kong, Frederic F Leymarie, Xavier Orriols, Eileen L Vote, Martha S Joukowsky, Benjamin B Kimia, David H Laidlaw, David Mumford, Senem Velipasalar. Assembling virtual pots from 3D measurements of their fragments. In: Proceedings of the VAST (Euroconference on Virtual Archeology), 2001.
[54] Victoria Interrante. Illustrating surface shape in volume data via principal direction-driven 3d line integral convolution. In: Proceedings of the ACM SIGGRAPH ’97, August 1997. pp. 109–116.
[55] Chris G Healey, J T Enns. Large datasets at a glance: combining textures and colors in scientific visualization. IEEE Transactions on Visualization and Computer Graphics 1999;5(2):145–167.
[56] David H Laidlaw. Loose, artistic ‘Textures’ for visualization. IEEE Computer Graphics and Applications 2001;6–9.
[57] David H Laidlaw, R Michael Kirby, J Scott Davidson, Timothy S Miller, Marco DaSilva, William H Warren, Michael Tarr. Quantitative comparative evaluation of 2D vector visualization methods. In: Proceedings of the IEEE Visualization ’01, October 2001.
[58] Joseph LaViola, Daniel Acevedo Feliz, Daniel Keefe, Robert Zeleznik. Hands-free multi-scale navigation in virtual environments. In: Proceedings of the ACMSIGGRAPH I3D 2001 (Symposium on Interactive 3D Graphics), March 2001.
[59] Glen Wheless, Cathy M Lascara, Donna Cox, Robert Patterson, Stuart Levy, Andrew Johnson, Jason Leigh, Ahbinov Kapoor. The use of collaborative virtual environments in the mine countermeasures mission. In: Proceedings of the SPIE’s 13th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls, April 1999.

[60] Jason Leigh, Andrew E Johnson, Thomas DeFanti, Stuart Bailey, Robert Grossman. A tele-immersive environment for collaborative exploratory analysis of massive data sets. In: Proceedings of the ASCI '99, June 1999. p. 3–9.

[61] Lew C Hill II, Carolina Cruz-Neira. Palmtop interaction methods for immersive projection technology systems. In: Proceedings of the PT 2000 (Immersive Projection Technology Workshop), Hosted by Virtual Reality Applications Center (VRAC), Iowa State University, Ames, IA, June 19–20, 2000.

[62] Andrew S Forsberg, David H Laidlaw, Andries van Dam, R Michael Kirby, George E Karniadakis, Jonathan L Elion. Immersive virtual reality for visualizing flow through an artery. In: Proceedings of the IEEE Visualization '00, October 2000.

[63] Tecplot home page: http:///www.amtec.com/Product_pages/tecplot9/index.html

[64] Victoria Interrante, Henry Fuchs, Stephen M Pizer. Conveying the 3D shape of smoothly curving transparent surfaces via texture. IEEE Transactions on Visualization and Computer Graphics 1997;3(2):98–117.

[65] David H Laidlaw, Eric T Ahrens, David Kremers, Matthew J Avalos, Carol Readhead, Russell E Jacobs. Visualizing diffusion tensor images of the mouse spinal cord. In: Proceedings of the IEEE Visualization '98, October 1998.

[66] NCSA Renaissance teams: http://www.ncsa.uiuc.edu/News/Access/Releases/98Releases/981117.IU-technologist.html; Donna Cox: http://www.ncsa.uiuc.edu/People/cos/vita.html

[67] Fred P Brooks, Jr. The mythical man-month. Reading, MA: Addison Wesley, 1995 [revised edition].

[68] Daniel F Keefe, Daniel Acevedo Feliz, Tomer Moscovich, David H Laidlaw, Joseph J LaViola. CavePainting: a fully immersive 3D artistic medium and interactive experience. In: Proceedings of the ACM SIGGRAPH 13d 2001 (Symposium on Interactive 3D Graphics), 2001.

[69] Marco J DaSilva, Song Zhang, Catagay Demiralp, David H Laidlaw. Visualizing the differences between diffusion tensor volume images. In: Proceedings of the International Society for Magnetic Resonance in Medicine Diffusion MRI Workshop, 2000. URL: http://www.ismrm.org/workshops/diffusion/index.html

[70] LLNL ROAM home page: http://www.llnl.gov/graphics/ROAM/

[71] Henry Fuchs, Mark Livingston, Gary Bishop, Greg Welch. US patent #5870136. Dynamic generation of imperceptible structured light for tracking and acquisition of three dimensional scene geometry and surface characteristics in interactive three dimensional computer graphics applications, February 9, 1999.

[72] Andres M Kunz, Christain P Spagno. Modified shutter glasses for projection and picture acquisition in virtual environments. In: Proceedings of the IEEE Virtual Reality, 2001. p. 281–282.

[73] Wojciech Matusik, Chris Buehler, Ramesh Raskar, Steven J Gortler, Leonard McMillan. Image-based visual hulls. In: Proceedings of the ACM SIGGRAPH, 2000. p. 369–374.

[74] Jane Mulligan, Kostas Daniilidis. View-independent scene acquisition for tele-presence. In: International Symposium on Augmented Reality, Munich, 2000.

[75] Greg Humphreys, Ian Buck, Matthew Eldridge, Pat Hanrahan. Distributed rendering for scalable displays. In: Proceedings of the ACM Conference on Supercomputing, 2000.

[76] Elumen's VisionStation: http://www.elumens.com/products/visionstation.html, and Fakespace ConCave: http://www.fakespacesystems.com/

[77] Naval Postgraduate School MOVES Institute Virtual Environment design: http://www.movesinstitute.org/Projects/MOVESresearchcenter.html

[78] USC's VET (Virtual Environment Training) institute: http://btl.usc.edu/VET/

[79] CMU's Stage Three design group: http://www.alice.org/stage3/

[80] University of Central Florida's CREAT Digital Media Program: http://www.creat.cas.ucf.edu/