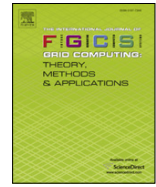




Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

The StarCAVE, a third-generation CAVE and virtual reality OptIPortal

Thomas A. DeFanti^{a,*}, Gregory Dawe^a, Daniel J. Sandin^b, Jurgen P. Schulze^a, Peter Otto^a, Javier Girado^c, Falko Kuester^a, Larry Smarr^a, Ramesh Rao^a

^a California Institute for Telecommunications and Information Technology (Calit2), University of California San Diego (UCSD), United States

^b Electronic Visualization Laboratory (EVL), University of Illinois at Chicago (UIC), United States

^c Qualcomm, Inc., United States

ARTICLE INFO

Article history:

Received 19 February 2008

Received in revised form

18 July 2008

Accepted 22 July 2008

Available online 17 August 2008

Keywords:

CAVE, Computer-supported collaborative work

Distributed/network graphics

Graphics processors

Graphics packages

Image displays

Interactive environments

Network communication

Sonification

Teleconferencing

Virtual reality

ABSTRACT

A room-sized, walk-in virtual reality (VR) display is to a typical computer screen what a supercomputer is to a laptop computer. It is a vastly more complex system to design, house, optimize, make usable, and maintain. 17 years of designing and implementing room-sized “CAVE” VR systems have led to significant new advances in visual and audio fidelity. CAVEs are a challenge to construct because their hundreds of constituent components are mostly adapted off-the-shelf technologies that were designed for other uses. The integration of these components and the building of certain critical custom parts like screens involve years of research and development for each new generation of CAVEs. The difficult issues and compromises achieved and deemed acceptable are of keen interest to the relatively small community of VR experimentalists, but also may be enlightening to a broader group of computer scientists not familiar with the barriers to implementing virtual reality and the technical reasons these barriers exist.

The StarCAVE, a 3rd-generation CAVE, is a 5-wall plus floor projected virtual reality room, operating at a combined resolution of ~68 million pixels, ~34 million pixels per eye, distributed over 15 rear-projected wall screens and 2 down-projected floor screens. The StarCAVE offers 20/40 vision in a fully horizontally enclosed space with a diameter of 3 m and height of 3.5 m. Its 15 wall screens are newly developed 1.3 m × 2 m non-depolarizing high-contrast rear-projection screens, stacked three high, with the bottom and top trapezoidal screens tilted inward by 15° to increase immersion, while reducing stereo ghosting. The non-depolarizing, wear-resistant floor screens are lit from overhead. Digital audio sonification is achieved using surround speakers and wave field synthesis, while user interaction is provided via a wand and multi-camera, wireless tracking system.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

A key criterion for VR is the provision of a “immersive” display with significant tracked stereo visuals produced in real time at larger angle of view than forward-looking human eyes can see. Immersion can be provided by head mounted displays and often is. Another means for immersion is the room-sized projection-based surround virtual reality (VR) system, variants of which have been in development since at least 1991 [1–3]. The first CAVE¹ prototype was built in 1991, showed full scale (3m³) in

public at SIGGRAPH'92² and SC'92, and then CAVEs were built for the National Center for Supercomputing Applications, Argonne National Laboratory, and The Defense Advanced Research Projects Agency. In the past 17 years, hundreds of CAVEs and variants have been built in many countries. Software called “CAVELib” [4] was developed and is still widely in use.

The first generation CAVE used active stereo (that is 96–160 fps field-sequential images separated by glasses that synchronously blink left and right) to maintain separate images for the left and right eyes. Three-tube cathode ray tube (CRT) Electrohome ECP and then Marquee projectors (with special low-persistence green phosphor tubes) were used, one per 3 m² screen, at a resolution of 1280 × 1024 @ 120 Hz, thus displaying about the equivalent of

* Corresponding author. Tel.: +1 312 996 3002; fax: +1 312 413 7585.

E-mail addresses: tdefanti@ucsd.edu, tom@uic.edu (T.A. DeFanti), gdawe@ucsd.edu (G. Dawe), dan@evl.uic.edu (D.J. Sandin), jschulze@ucsd.edu (J.P. Schulze), potto@ucsd.edu (P. Otto), jgirado@gmail.com (J. Girado), fkuester@ucsd.edu (F. Kuester), lsmarr@ucsd.edu (L. Smarr), r Rao@ucsd.edu (R. Rao).

¹ The name CAVETM was coined by the lead author of this paper for the VR room being built at the Electronic Visualization Laboratory (EVL), University of Illinois

at Chicago (UIC), which was subsequently commercialized by the company that is now Mechdyne, Corporation.

² Michael Deering of Sun Microsystems, Inc. exhibited a 3-wall similar system for one user called the *Portal* at SIGGRAPH'92 [3]. The 3-wall+floor CAVE at SIGGRAPH'92 allowed multiple users.

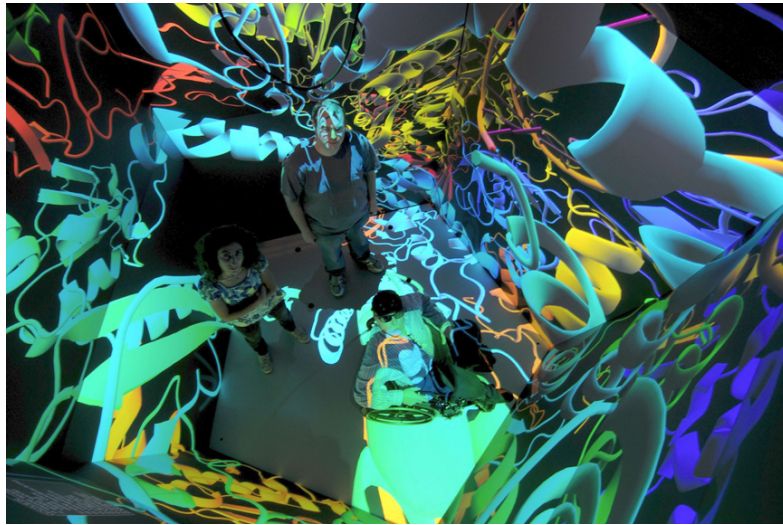


Fig. 1. The StarCAVE from above, looking down on a RNA protein rendering. The still camera taking the picture is not being tracked so the perspective is skewed, but this image shows the floor as well as the walls and shows some of the effects of vignetting and abnormally severe off-axis viewing.

20/140 to 20/200 visual acuity.³ Besides providing an experience of “low vision”,⁴ the first CAVEs were relatively dim (the effect was like seeing color in bright moonlight⁵), and somewhat stuttering (the networked Silicon Graphics, Inc. (SGI) workstations, one per projector, could maintain only about 8 updates of a very modest 3-D perspective scene per second, insufficient for smooth animation). Ascension, Inc. Flock of Birds electromagnetic tethered trackers were used to poll the 6 degree-of-freedom (DOF) position of the user’s head and hand. There were three rear-projected walls and a down-projected floor, which gave a then novel complete feeling of room-sized immersion. The screen frame was made of non-magnetic steel to decrease interference with the tracker, and the screen was a grey flexible membrane screen stretched over cables in 2 corners. About 85% of the cost of the first generation CAVE was in the 5 SGI Crimson workstations, later the 4-output 8-processor SGI Onyx.

A second-generation CAVE was developed by EVL in 2001,⁶ featuring Christie Mirage DLP 1280 × 1024 projectors that are 7 times brighter⁷ than the Electrohomes of the first generation, although 5 times the cost. Users’ color perception got much better because of the brighter projectors delivering adequate light to their eyes’ color receptors. Since chip-based projectors (LCD, LCOS, DLP) do not have the numerous analog controls on sizing that the CRT projectors did, there is no available electronic adjustment on modern projectors for keystone and other distortions. Mechanical optical alignment requires precision of frame fabrication, projector mounts, and the flatness and squareness of the projector image through the lens to achieve

accuracies of 1 pixel in 1000. Now that there was brighter projection, the screen material also had to be chosen carefully to maximize contrast and minimize internal light spillage on the other screens (too much of which reduces the contrast of the images and makes them look washed out).⁸ (None of these problems occur with normal use of projectors since they are not typically edge-matched, especially on multiple edges.) This system also used active stereo at 60 Hz/eye (the projectors update at 120 Hz) and could, with the SGI Reality Engine, get ~25 graphic scene updates per second, a 3x improvement over the first-generation SGI Crimsons, resulting in much smoother motion. For this CAVE, about 60% of the cost was in the SGI 8-processor shared-memory cluster. This DLP-based CAVE is still built⁹ and sold, although PCs now run the 1280 × 1024 screens (often cropped to 1024 × 1024). The acuity is still roughly 20/140 acuity from the center of a 3 m CAVE that has ~1 megapixel/screen.

EVL’s research focus has always been aimed at practitioners of scientific visualization and artists. While the first and second generation CAVEs were quite effective in conveying immersion, the 20/140 visual acuity resulted in “legally-blind” scientific visualization, admittedly contradiction in terms, but it was state-of-the-art VR at the time nonetheless. The number of pixels per screen was impractical to improve with the projector technology of the time, so instead EVL research started to focus on tiled displays with dozens to hundreds of megapixels [6]. By adopting what we learned from building these tiled displays and the computer clusters that drive them, we were able to design the StarCAVE, a third-generation tiled CAVE completed in July 2007 at Calit2 at the University of California in San Diego; the down-projected floor was added in May 2008. (See Figs. 1–3) The StarCAVE exploits tiled visual parallelism to increase visual acuity to ~20/40¹⁰ from 3 m away, and brightness to ~6 foot Lamberts (6fL), through

³ Calculated by assuming 100 pixels/ft from 10’ away giving 6.84 arc min/pixel, or ~20/137. At this quality in vehicle interior simulations, for example, the odometer/speedometer gauges can’t be read, even when in focus (as everything in a VR scene typically is). The metric equivalent of 20/20 is 6/6; 20/200 is 6/60. Randy Smith of GM Research says, in an unpublished technical report, that GM’s 2.5 m² CAVE’s acuity is (was) 20/200 [5].

⁴ In typical drivers’ license exams, anything worse than 20/40 requires additional evaluation; 20/200 in the worse eye is considered legally blind. *Low vision* is sometimes used to describe visual acuities from 20/70 to 20/200, according to <http://en.wikipedia.org/wiki/Blindness>.

⁵ See details in http://flywheel.caset.buffalo.edu/wiki/Image:Dcp_2640.jpg.

⁶ See <http://www.evl.uic.edu/pape/CAVE/DLP/>.

⁷ The Christie Mirages claim 5000ANSI lumens, which on a 3 m × 3 m screen of this type, through the active stereo eyewear, yields about 5fL brightness.

⁸ We used a “Disney Black” screen, its unofficial name, from Stewart (<http://www.stewartfilm.com/>). It has never been clear what the official name is, but one can ask for it unofficially and get it.

⁹ See <http://www.fakespace.com/cave.htm>.

¹⁰ We used a scanned-in “illiterate” eye chart (the one with the E’s in various directions) at the proper viewing distance to judge the acuity. We assume the projectors are at optimal focus. 20/40 was subjectively discernable from 3 m. At 1.5 m from the screen, we discerned 20/60.

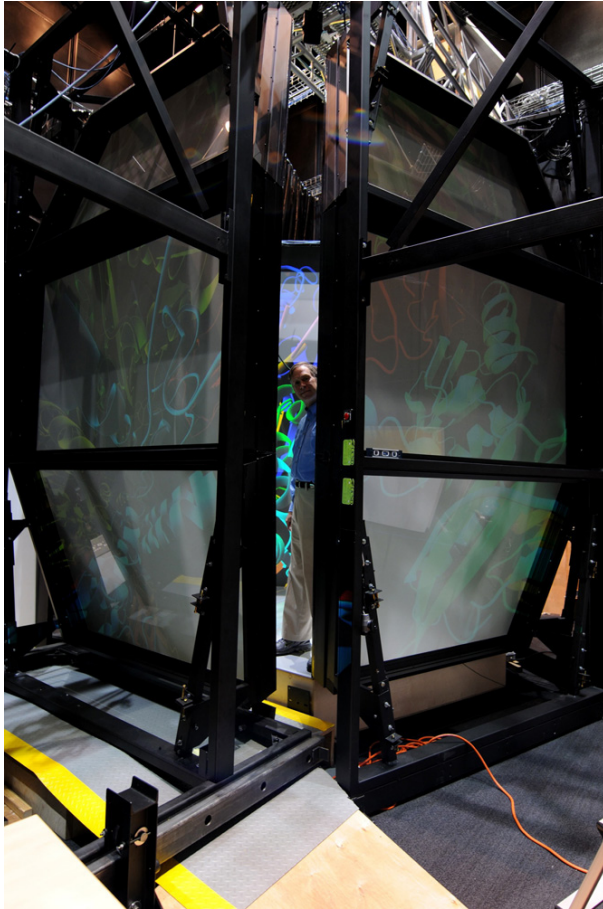


Fig. 2. A photograph of the StarCAVE from the outside showing the entry door nearly closed.

the glasses,¹¹ uses LCOS LCD projectors with integrated circular polarization coupled with passive lightweight circularly polarized eyewear similar to sunglasses, instead of much heavier battery-powered active shutter glasses used with second generation CAVE DLP projectors. The StarCAVE achieves approximately 34 megapixels per eye, fully surround, including the floor (but no ceiling). The cost is about the same as the first and second generation CAVEs, except in 2007, the computers were only 10% of the cost, the projectors 40%, the screens and frames 40% and the tracking and audio the remaining 10%.¹² The StarCAVE was primarily funded by Calit2, and is also supported by the OptIPuter Project.¹³

Also deserving of note is what we describe as a fourth-generation CAVE, called the Varrier™,¹⁴ developed at EVL [7] and replicated at Calit2. It is a 40 megapixel per eye autostereo device (meaning no stereo glasses are worn by the user), and it is made out of 60 LCD 1600 × 1200 panels (not projectors). It is also part of the OptIPuter Project. See also EVL's work on Dynallax [8], which is an outgrowth of Varrier.

¹¹ Commercial movie theater screens typically achieve 14fL brightness. Brightness on stereo movie theater screens is not standardized, but ranges from 0.5 to 7fL. The metric equivalent of 1 fL is 3.426 candelas/m².

¹² Another 3rd generation CAVE-like system, the C-6, can be seen at Iowa State, see <http://www.public.iastate.edu/~nscentral/news/2007/mar/C6.shtml>; It uses 24 Sony SXR4 4K projectors to get 100 megapixels per eye spread out on six screens.

¹³ NSF Cooperative agreement ANI-0225642.

¹⁴ Varrier is a trademark of the Board of Trustees of the University of Illinois.

2. Design of the StarCAVE

The StarCAVE was designed to be replicable by committed researchers. In summary form, additional design criterion were:

- Fully-tracked VR immersion (360° look-around capability) for 1 prime user and several observers
- bright, high-resolution, high-contrast visuals at ~ 20/40 vision
- lifelike, immersive ambient auditory experience, dynamically scaled to arbitrary virtual scenes, for one or more listeners, which necessitates a platform to explore and evaluate
 - multiple approaches to immersive audio experience, and
 - sonification of scientific data in certain visualization contexts
- straightforward acquisition, operation, maintenance, and possible upgrades with commodity projectors and computers.

Within the following constraints:

- ADA compliance (wheelchair accessibility)
- a low-as-possible ambient noise environment
- seismic safety compliance
- a substantial but not unlimited budget (US\$1,000,000)
- non-depolarizing rear and front projector screens.

The StarCAVE is contained in a room that is approximately 9.15 m³, which affected the choice of projectors, projection distances, screen sizes, cable lengths, and computers.¹⁵ ADA compliance was also a consideration that made installing a rear-projection floor (that is, projecting from the underneath) for us impossible—there is no practical and safe way to get a wheelchair onto a raised floor screen 3 m up in the air unless one installs an elevator, not an option in an existing space.¹⁶ Therefore, we decided on a down-projection reflective floor, and rear-projection horizontal surround screens, and no ceiling, a compromise that has worked acceptably in the past with earlier CAVEs, and works even better in the StarCAVE. The floor is painted with marine-quality polyurethane that has aluminum particles suspended in it, a commercially available product that, unlike normal front-projection polarizing screens, can be walked on.

360° surround VR changes the viewing paradigm so that the user never needs to horizontally rotate the image as one must when the screens fail to go all the way around, but one does need to design a means for entry. We created a door by putting one of the 5 walls on tracks perpendicular to that wall, which allow it to slide open and closed (Fig. 3).

Building a CAVE-like display with passive polarization on more than one screen, instead of using active shutter glasses, presents challenges perhaps obvious only to a practitioner of projected virtual reality, mainly due to the complexity of maintaining polarization with rear projection, in particular. A polarized projector beam cannot not be folded using mirrors at angles more than ~40°, lest the polarization be lost, an issue because folding is normally used to minimize the space needed outside a CAVE.¹⁷ The rear projection screens need to be polarizing-preserving, not normally the case with rear-projection screens, and as diffuse (low gain) as possible to minimize hot-spotting (see Fig. 4). The spec

¹⁵ We recommend to anyone building a CAVE to get a bigger room. 6-wall CAVE-like systems in Stockholm, Gifu, Iowa, and Louisiana have used mirrors or had big enough rooms. Of course, one could request from the architects a 3 m² hole in the floor and ceiling of a 3-story space, but this is harder done than said, in practice.

¹⁶ We actually did specify an elevator, and one was provided, but the pathway to the StarCAVE space was unfortunately blocked by a huge insurmountable steel beam added during the value engineering phase, not to our knowledge until too late.

¹⁷ Normal projector lenses require a throw distance of at least 2x the screen width, and there needs to be room for the projector body as well. This means a 3 m² CAVE with normal lensing and unfolded optical paths requires a room ~18 m on the diagonal, minimally, with no obstructions. Not having a space that big or the option of folding the optics, we procured, instead, very wide angle lenses that require only 1x the screen width, and smaller screens.



Fig. 3. A photograph taken with the camera being tracked of the simulated interior of the Calit2 Building at UCSD. Note the minimal seams, excellent perspective rendering, and the effect of the floor. Effects of vignetting and off-axis viewing are seen here as well, far more noticeable in still photographs than perceived when experienced live.



Fig. 4. Hotspotting of illumination (left) versus more even dispersion (right) on polarizing preserving screens with different coatings.

sheets on available screen materials describe polarizing preserving attributes in optimistic qualitative, not quantitative terms; all we tested failed to meet our requirements of less than 3% ghosting on center. We worked with a custom screen manufacturer¹⁸ for several months to iteratively develop and test coatings, and created a rigid screen with excellent characteristics as quantitatively measured and qualitatively perceived. Thus, we use screens that are 1.2 m by 2.13 m coated PMMA (polymethyl-methacrylate) rigid plastic, illuminated from the back by JCV HD2K projectors with 1:1 wide-angle lenses. All the projectors and screens need to be held in place to sub-pixel precision, so a steel exoskeleton

¹⁸ The single element rigid grey screen we use is the custom-fabricated rps Visual Solutions' "rps ACP Cinepro Ultra Screen". It has a specialized coating on a substrate acrylic which creates an opaque screen. rps/ACP were extremely generous with their time and effort in reformulating screen coatings until we were able to achieve polarization separation of better than 5 stops (~98%) in the middle and 3.4 stops (~90%) at the screen corner from 1.5 m away. This screen has a remarkable contrast ratio of ~8.5:1 and transmits about 50% of the rear-projected light. We needed to use rigid screens because the typical flexible membrane screens used in CAVES billow with air pressure, and in our case, would likely sag on the tilted top and bottom rows, plus it is not known how to effectively coat a flexible screen with this polarizing-preserving high-contrast spray coating.

and screen corner details were designed by Calit2's Greg Dawe, rpVisual Solutions, Inc. and ADF, Inc. and fabricated by computer-assisted means. The screens are positioned in 5 columns of 3 panels high, with the top and bottom ones tilted in by 15° to maximize the practical on-axis user view of all the projectors¹⁹ (see Fig. 5). This required manufacturing and supporting 10 trapezoidal and 5 rectangular screens to very fine mechanical and structural tolerances (0.1 mm, see Fig. 6), considering the size of the StarCAVE. As noted above, one of the 5 columns, along with its 6 projectors, 3 screens, and 3 computers, rolls in and out 1.5 m on tubular steel rails, thus providing access for people to the interior of the StarCAVE.

The trapezoidal/pentagonal shape, an optimized solution to fitting in the 9.144 m² physical room, also turns out to have many advantages over the standard cubic CAVE. Interior light spillage has been noticeable with cubic CAVES, especially in variants that did not use contrast-preserving (dark) screens. Since they form a pentagon, none of the StarCAVE matte-surfaced screens directly reflects on any other into the user's eyes. The 108° angle between the screens (instead of 90°) means less light spillage from screen to screen, and screen to floor (a 105° angle). The viewer also sees somewhat less oblique views in that the typical comfortable viewing angle is less off-axis than in a standard CAVE because there are 5 screens²⁰ instead of 4. The tilted-in top screens also mean that the lack of a ceiling screen is not much noticed—one really has to look uncomfortably straight up to run out of projected image.²¹ In the Future Research section, we discuss user-centric ghost and vignette mitigation, which can help improve the stereo separation and further normalize the illumination in the StarCAVE and conventional CAVES.

¹⁹ 15° was a carefully chosen angle; a few degrees more than 15 would be more on-axis to the user, but the floor size, as well as the ceiling hole, would be diminished by the increased angle. The floor-screen height is limited to a two step rise above the room floor to allow sufficient room for Americans with Disabilities Act (ADA)-compliant ramps within the room. The alternatives, demolition of the concrete floor and excavation or the procurement and installation of an ADA compliant lift and attendant ramps or demolition, were judged too costly.

²⁰ 6 and 7 screen columns were also considered but wouldn't fit in the room unless the screens were shrunk accordingly, which would make 3 of them not high enough for standing people.

²¹ 6-screen cubic CAVES, when the entry is tightly closed, have potential air supply considerations. A CAVE full of people may well quickly up all the oxygen but this has never been experimentally verified to our knowledge.

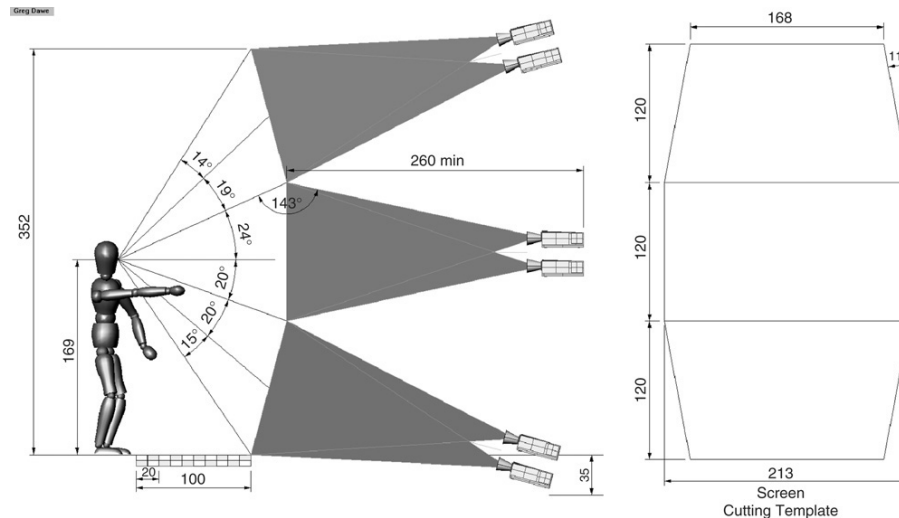


Fig. 5. StarCAVE diagram showing sectional field of view, projector light paths, and screens (measurements in inches).

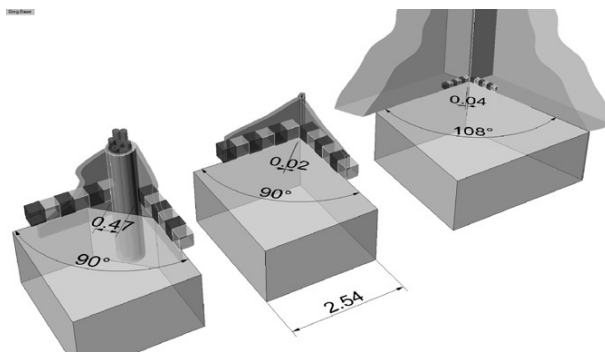


Fig. 6. Detail showing size of pixels and number of which are blocked by screen edge supports in (left to right) 1st, 2nd, and 3rd generation CAVEs, in inches.

3. Choice of the StarCAVE projectors

As mentioned above, we chose to use polarized stereo for the StarCAVE.²² Achieving stereo in any projected VR system typically absorbs $\sim 90\%$ of the projector lumens because one needs to polarize both the projector and the user's glasses, losing between 3 and 4 stops (88%–94%) of the light. In order to minimize internal light spillage and maintain contrast, rather dark grey screens are used in successful CAVEs, which means another stop of light is typically lost (5 stops is 97%, which means only 3% of the light gets through to the screen). One method to get enough light is to use theater-bright (5000–10,000 lumen) projector(s), but these are $\sim \$100,000$ each, take significant power (2kW for 5000 lumens and 4kW for 10,000 lumens) and cooling, and they are physically big and heavy. Fortunately, JVC makes HD2K projectors which are designed for stereo simulator installations, and which output cleanly polarized light. By reversing the direction of the circular polarization with a retarder plate, with near zero loss of light, one polarizing filter is eliminated, gaining about a stop in brightness. These $1920 \times 1080p$ projectors are about 1/10th the weight, cost and power of big theater projectors and accept short throw wide-angle lenses, which means the throw distance is the same

as the screen width, a critical feature, as noted above, for fitting in a room without folding the projection with mirrors (which, remember, will de-polarize the light if at an angle $>40^\circ$). Since the HD2Ks are projecting on screens that are about 1/4th the size of classic 3 m^2 CAVE screens, another two stops of brightness are gained, so even with these claimed 800-lumen projectors (which we measure to be 300 lumens after 100 h of lamp use) the StarCAVE affords the user 6fL/eye through the circularly polarized eyewear at 50–1 contrast.²³ JVC provided us software to intensity- and gamma-match the 34 projectors so they are all adjusted to the same color and brightness (within $\pm 1/5$ stop), an intractable manual task. We did, however, need to manually position and align the projectors, which took days of work and is not perfect. (See the Future Research section for thoughts on auto-alignment techniques.) One feature is that these projectors can be remotely controlled through their serial ports, which means one does not need climb scaffolding to manually turn them on and off (which would really have been manual since we had to tape over the projector's remote control receivers because the tracker camera infrared illuminators were interfering with them!).

4. Choice of StarCAVE computers

Every projector pair for the 15 vertical screens and two half-floor screens is driven by a computer node (specifically, an Intel quad-core Dell XPS 710 computer using ROCKS [10] and CentOS [11], with dual Nvidia Quadro 5600 graphics cards). We use an additional XPS 710 machine as the head node to control the rendering nodes, for a total of 18 nodes. The nodes are interconnected with 1 Gigabit Ethernet and 10 Gigabit Myrinet networks and connected to out-of-room networks with 5×10 Gigabit Ethernet. We chose these particular Dells because they were available with the 1KW power supplies needed to power the dual 5600 cards and because they have large ($\sim 12 \text{ cm}$) fans

²² We also investigated color-separation stereo encoding with Infitec [9] filters, but this technique only works well with xenon lamps which only are found on theater-class projectors, well out of our price and power budget.

²³ Details: readings taken with Minolta F Spotmeter at ISO setting of 100 while multiple screens displayed 4×4 black & white pattern from the center of the StarCAVE. Projector output at the (projector side) center of the screen: $\sim 12\text{fL}$. White field at the (viewer side) center of the screen: $\sim 8.5\text{fL}$ ($\sim 1/2$ stop loss due to screen). White (through eyewear) at center screen: 6fL ($\sim 1/2$ stop loss due to eyewear). Black (through eyewear) at center screen: less than $1/10\text{fL}$, thus $>50:1$ contrast ratio. Worst case white through eyewear) off axis from StarCAVE center to a screen corner: $\sim 1.3\text{fL}$, which is 2 stops, a significant amount.

(unlike rack-mounted PCs) that are relatively quiet. (Low fan noise is important in any audio-visual environment if quality audio using speakers is a goal. We could not move the computers to another room because DVI copper cables for $1920 \times 1080p$ images are (maximally) about 15 m long; extending the 34 dual-link DVI signals over fiber to our server room would cost about as much as the computers themselves, so we carefully listened to the computers fan noise before we bought them!²⁴) We also bought spares of parts not easily replaceable: one rectangular and one trapezoidal screen, 2 XPS 710/5600 computers, and four HD2K projectors which we use for a 2-screen development system in a VR lab, but which are available for StarCAVE maintenance when needed. Both the specific computers and projectors we use have already been discontinued, replaced by newer models, so it is considered critical for maintenance to have identical spares bought with the rest and set aside.

5. Tracking

For head and hand tracking we use a wireless, optical tracking system by ART Tracking [12]. It consists of four infrared cameras and a Flystick2, a six degree-of-freedom (DOF) device with five buttons and an analog joystick. We mounted the cameras at the top of the walls. The ART system communicates via Ethernet with a VRCO Trackd [13] demon running on the head node of the rendering system. Trackd provides the application developer with 6 DOF positions for hand and head, as well as the status of all buttons and the joystick on the joystick. The ART system has an accuracy of under a millimeter at 60 Hz.

6. Audio for the StarCAVE

Two very significant impediments are encountered when trying to provide an auditory experience to match the resolution and immersive quality of the StarCAVE's visuals:

- The StarCAVE's virtually seamless, 360-degree rear projection screen construction precludes the use of laterally placed loudspeakers which are normally used to enhance spatial imaging of sound. Lateral speakers would cause shadowing or block the user's view of screens.
- The acoustical environment is extremely difficult to control given the StarCAVE's extremely acoustically reflective screens in a symmetrical deployment, devoid of free-field or absorptive attenuation, which causes unusual modal artifacts (e.g., resonances and cancellations). These considerations shaped all ensuing decisions concerning the physical design of the system.

Much discussion focused on driver element designs and materials. Ribbon tweeters, planar speakers and thin line arrays were investigated but rejected as too difficult to mount or control. Head-tracked, binaural, headphone approaches were identified and explored as an obvious choice for a single user solution, and an implementation using new, custom software is under development. However, headphone-based, head-tracked systems are unsuitable for multiple simultaneous users who may also want to be conversing with each other. Since space and therefore speaker size was limited, a sub-satellite speaker system, as found in consumer and professional 5.1 systems, was chosen.

We investigated new small speaker enclosures designed by MeyerSound, Inc. for use in the LCS Constellation active reverberation enhancement system. These small, wide-range, wide-dispersion speakers, measuring 10 cm^2 , were selected for

audio quality, robustness, positional versatility, their integrated equalization and system controls, and enthusiastic manufacturer support. It was hypothesized that a combination of direct and reflected sound from three arrays of 5 speakers would provide ample coverage for either 5.1 multichannel audio (which matches well to the 5-sided StarCAVE environment) or up to 15 channels of discreet audio diffusion. Models of radiation and absorption patterns were sketched onto design drawings and a plan for exact speaker placement evolved:

- *Cluster 1:* 5 MM4 speakers mounted in a tight cluster overhead with radiation patterns directed at the upper portion of the 5 screens, which would in turn reflect at a lowered angle onto the ears of the listener.
- *Cluster 2:* 5 MM4 speakers mounted at the top-corner of each upper screen, sharing the mounting structure of the infrared positional tracking emitters, radiating directly.
- *Cluster 3:* 5 MM4 speakers mounted at the bottom-center of each lower screen, angularly directed at the ears and exiting through the open top of the structure
- A subwoofer channel was designed into the flooring structure of the StarCAVE, with ports in each of the five corners.

15 channels of amplification are provided by QSC Multichannel power amps. Equalization and panning are controlled by 2 Meyer/LCS Matrix 3 units, calibrated by Richard Bugg of Meyer-Sound. Additional analysis equipment was brought in to help with system calibration. Extensive equalization was required, as analysis revealed extensive spectral non-linearities, as anticipated (see Fig. 7).

Run-time electronics and panning was initially handled by a large MaxMSP [15] patch running ICST Ambisonics [14] and using two MOTU multichannel interfaces. All source and control, with the exception of system EQ, has now been integrated into a series of MaxMSP patches running on a Mac desktop computer. Soundfile playback is handled here along with tone and noise source testing. Several approaches to panning and imaging are implemented including Ambisonics, VBAP [16] and UCSD's Designer-derived spatializers [17]. Positional and user input is tracked from a StarCAVE Linux computer, and communicated to the audio host.

Next, we installed a Yamaha YSP-1100 sound projector, which uses 40 speakers in a single rectangular enclosure $1,030 \times 194 \times 118 \text{ mm}$ in size²⁵ (see Fig. 8). This approach uses a derivative of wave-field synthesis [18] (WFS) to direct tightly focused "beams" of audio, in order to provide consumers with surround sound without using rear speakers. The system depends upon bouncing first reflections off the side walls; in this context the extreme acoustic reflectivity of the StarCAVE actually worked in our favor. The Yamaha YSP technology was difficult and time-consuming to calibrate and several positions were tried before we began to achieve some degree of convincing surround sound. The YSP device will only process Dolby or DTS encoded surround audio, which limits its immediate usefulness without modification or additional equipment.

The initial MM4 implementation produced reasonably good results with some sense of immersion and good audio fidelity. Users noticed that loud levels overloaded the space, resulting in degraded intelligibility, and a sense of missing lateral energy was noticed, as was an impression of too much overhead energy. The imagery was somewhat unstable and lacking in spatial clarity, which we hypothesize is due to the wide diffusion patterns of the speakers in a too-reflective environment.

²⁴ Noise level of PCs is not something given in manufacturers' specifications, nor is this information available on the web.

²⁵ See http://www.yamaha.com/yec/ysp1/idx_spec_ysp1100.htm.

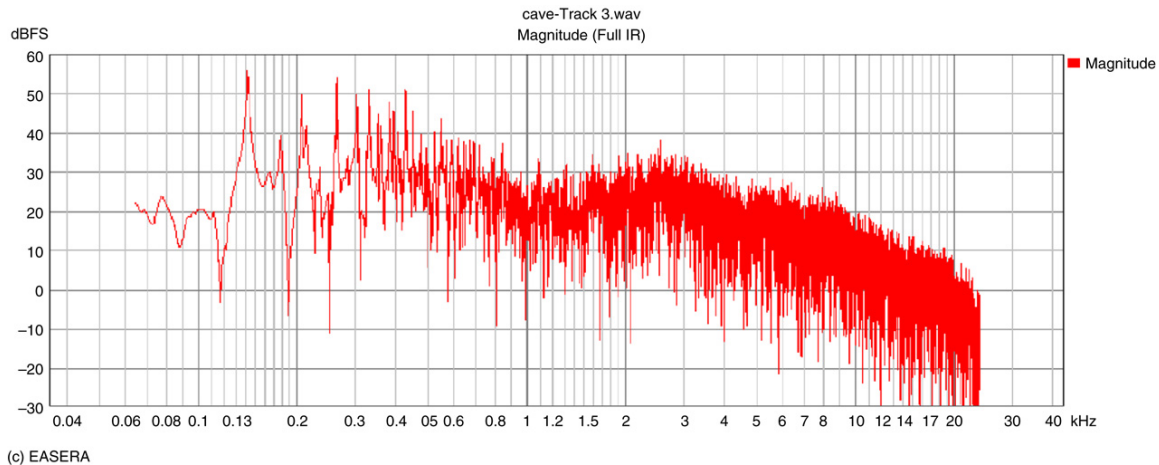


Fig. 7. Full range impulse response of StarCAVE before equalization, showing extreme spectral non-linearities.

Performance of the floor-mounted subwoofer was found to be sufficient for its intended use, though low frequencies are noticeably localized in the lower part of the structure. An additional overhead subwoofer should alleviate that problem. Listeners respond favorably to the low-frequency physical sensation of bass transferred through the floor.

Comparison of the performance of the various spatial processing and diffusion approaches is a complex task and the subject of ongoing study. Testing for subjective and objective data is being designed.

We found that although difficult and time-consuming to calibrate, the Yamaha YSP technology provided stable imagery and a good sense of immersion with somewhat less unintended overhead directionality than the initial approach. Fidelity was worse than the Meyer components, but the YSP has a wide variety of useful calibration features such as beam width and focal length, wide ranging horizontal and vertical directivity controls, and is relatively inexpensive. Image control was better than the Meyer components and the usefulness of the YSP is leading to the design and implementation of a custom WFS system.

7. Software for the StarCAVE

We currently support two software environments to drive the StarCAVE: OpenCover, which is the OpenSceneGraph-based VR renderer of COVISE [19], and EVL's Electro [20]. However, any software capable of driving a multi-screen and multi-node environment should be able to run in the StarCAVE (e.g., CAVElib), as long as it can accommodate the tilt of the upper and lower displays, as well as handle the pentagonal arrangement of the walls, both of which are, of course, non-standard configurations for immersive virtual reality systems. We use ROCKS-based OS distribution and management to quickly install and recover nodes. As mentioned before, the projectors are connected via their serial ports to simplify remote control of all their programmable features, including on/off.

8. Applications in the StarCAVE

Most of the applications that run in the StarCAVE were developed over the two years that our prototype 4-projector, two screen test version existed before the StarCAVE was finished. Once the StarCAVE's walls were configured in OpenCOVER [21], porting the software to the StarCAVE was mostly a matter of recompiling the source code. Since there are 17 rendering PCs,

we duplicate critical data on the 500 GB local drives on each PC, which considerably cuts down on startup time, if warranted. We are currently experimenting with distribution of data over the 10G Myrinet connections to each PC, since speed in loading models is important to users.

Among the new applications we wrote for the StarCAVE in 2007 and early 2008 are:

- The Protein Viewer (Fig. 1): This application connects the StarCAVE to the PDB data bank server [22] to download one of ~50,000 protein structures and display it in 3D. We use PyMOL [23] to convert the PDB files to the VRML format, which our visualization software can read. The user can choose between different visualization modes (cartoon, surface, sticks), load multiple proteins, align proteins, or display the corresponding amino acid sequence. The high resolution of the StarCAVE and its surround display capability allows the user to view a large number of protein structures at once, which the scientists do to find similarities between proteins.
- Bay Bridge: We created an application that displays CAD models of parts of the new San Francisco Oakland Bay Bridge in the StarCAVE. This application allows users to walk/fly through these parts at their real size to find material clashes, construction errors, and draw conclusions with regards to constructability.
- Neuroscience: We created a virtual replica of the Calit2 building (Atkinson Hall) to allow users to walk through it in the StarCAVE. We are currently working with neuroscientists to find out if the human brain works similarly for wayfinding tasks in the StarCAVE and in the real world (see Fig. 9.).
- OssimPlanet: We integrated the Google Earth-like software library OssimPlanet [24] with OpenCOVER, allowing us to fly over the earth with highly detailed satellite images and 3D terrain. We have used this API in different application domains, from the visualization of ocean currents to genomics.
- Palazzo Vecchio: We visualized a LIDAR scan of the Hall of the 500 with about 26 million points as a point cloud in the StarCAVE. Our custom data distribution and rendering algorithm achieves a rendering rate of about 10 frames per second.

9. Future work on the StarCAVE

The StarCAVE is a computer science research project as well as a science, engineering and art production facility. We would,

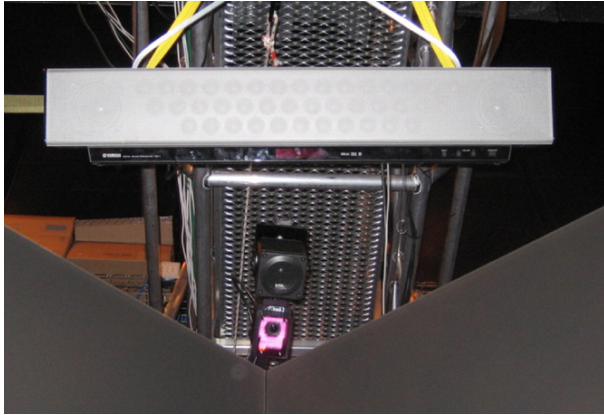


Fig. 8. Yamaha YSP-1100 sound projector (top), MM4 speaker (middle), and Art Tracker camera (bottom).



Fig. 9. Subject in the neuroscience project navigating the Calit2 building, Atkinson Hall. The stereo was not turned off for this photograph, hence the double image.

for instance, like to build on work by Aditi [25] and incorporate user locality-based spatial illumination compensation, since we have accurate 6 DOF information from the tracker. We can mitigate the vignetting due to projector angle and screen hotspotting by creating an inverse hotspot that tracks the user and is even dependent on the particular eye. This technique is possible in virtual reality environments today by taking advantage of the programmability of modern graphics hardware. A fragment shader can be written to take as an input the position of the hotspot and a brightness falloff value, and then this can be applied to every frame rendered by the system. In our head-tracked system, the location of the hot spot needs to be updated for every frame because one's head tends to be on the move, even if only slightly. This approach works only for one viewer, the head-tracked user. In order to support multiple users, the system would need to be modified to support multiple head-tracked users, and the display of multiple stereo image pairs on the screen [26]. In order for the algorithm to work, empirical measurements with a camera or light meter and good approximate 3D positions of lenses, screen, and viewer would be necessary.

In addition, for all but stark black situations, one could subtract out a small percentage of the other eye's image to mitigate the ghosting, again weighted by the user's tracked position since there

is more ghosting as one gets more off axis to the screen and/or projector. Since full mitigation would consume perhaps 2 stops of brightness, making the image 4 times dimmer, partial mitigation at least is worth investigating, especially if it only marginally degrades real-time performance. Several publications have been published on this topic, for the general case [27] as well as the application to passive stereo systems [28]. More research needs to be done to incorporate angular effects of the glasses (including the orientation to screen when circular polarization is used, the angle the glasses are held relative to screen, and the angle the viewing ray has to the screen).

Another area of research is auto-alignment, something that is well understood in software and available in firmware on high-end projectors, but would need to be implemented in the Nvidia 5600 GPU for the StarCAVE (given the HD2K projectors), in addition to the hotspot and ghost mitigation, both of which would also use the GPUs.

As noted above, the StarCAVE operates at about 20/40 visual acuity. In order to improve this to 20/20, we would need to upgrade from 1920 × 1080p projectors to 4K projectors (3840 or 4096 by 2160 or 2400), if their form factor were about the same. However, 4K projectors are likely to be bigger and cost ~\$100,000 each for a while, and we'd need to buy at least another 15 computers and 30 GPUs, find a projector manufacturer to internally polarize the red, green, and blue all the same direction (not normally the case), or switch to Infitec, and we'd need to provide up to 4x the power and cooling. Doubling the visual acuity would be a \$4,000,000 equipment upgrade at current prices, but prices fortunately do tend to drop over time.

As to future audio work, the StarCAVE head node can output spatialized sound by sending information about the sound sources to a Mac G5 computer running a Pd (Pure Data) [29] patch. Content generation software based in Max and Pd is frequent in our Calit2 environment, but have yet to be fully adapted to the StarCAVE systems. Similarly all components needed for convincing wireless binaural audio have been developed, but integration and refinement work remains. With the help of industry affiliate MeyerSound, a small group of dedicated computer music and engineering students are exploring and developing StarCAVE audio technologies.²⁶

Acknowledgements

UCSD receives major funding and from the State of California for the StarCAVE. UCSD receives funding for StarCAVE application development from the Gordon and Betty Moore Foundation for CAMERA, the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis. UIC has received numerous NSF, DARPA, and DOE awards for virtual reality research and development from 1992 to the present. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the funding agencies and companies.

²⁶ Audio assistance: MeyerSound: Richard Bugg, Steve Ellison, Will Lewis; Barry Threw, Jann Wissmuller, UCSD: Tom Erbe, Jann Wissmuller, Grace Leslie, Suketu Kamdar, Kevin Larke, Kyle Knickerbocker, Toshiro Yamada, Michelle Daniels and Chris Schlyer.

References

- [1] T.A. DeFanti, D. Sandin, C. Cruz-Neira, A 'Room' with a 'View', *IEEE Spectrum* (1993) 30–33, 10/01/.
- [2] C. Cruz-Neira, D. Sandin, T. DeFanti, R. Kenyon, J. Hart, The CAVE®: Audio visual experience automatic virtual environment, *Communications of the ACM* (1992) 06/01/.
- [3] M. Deering, High resolution virtual reality, *ACM SIGGRAPH Computer Graphics* 26 (2) (1992) 195–202.
- [4] <http://www.vrco.com/CAVELib/OverviewCAVELib.html>.
- [5] L. Baitch, R.C. Smith, Physiological correlates of spatial perceptual discordance in a virtual environment, in: *Proceedings Fourth International Projection Technology Workshop*, Ames, Iowa, 2000.
- [6] www.evl.uic.edu/cavern/lambdavisioin.
- [7] Daniel J. Sandin, Todd Margolis, Jinghua Ge, Javier Girado, Tom Peterka, Thomas A. DeFanti, The Varrier autostereoscopic virtual reality display, *ACM Transactions on Graphics: Proceedings of ACM SIGGRAPH 2005*, Los Angeles, CA, July 31-August 4, 2005, Vol. 24, No. 3, pp. 894–903.
- [8] T. Peterka, R.L. Kooima, J.I. Girado, Jinghua Ge, D.J. Sandin, A. Johnson, Leigh, J. Schulze, T.A. DeFanti, Dynallax: Solid state dynamic parallax barrier autostereoscopic VR display, *virtual reality conference, 2007 VR '07*, 10–14 March 2007, pp. 155–162.
- [9] <http://www.barco.com/VirtualReality/en/stereoscopic/infitec.asp>.
- [10] Philip M. Papadopoulos, Mason J. Katz, Greg Bruno, NPACI Rocks: Tools and techniques for easily deploying manageable Linux clusters, *Concurrency and Computation: Practice and Experience*, 15, (7–8), 707–725, In: Mark Baker, Daniel S. Katz (Eds.), *Cluster 2001* (special issue).
- [11] <http://www.centos.org/>.
- [12] <http://ar-tracking.eu/index.php>.
- [13] <http://www.vrco.com/trackd/Overviewtrackd.html>.
- [14] <http://www.icst.net/>.
- [15] <http://www.cycling74.com/products/maxmsp>.
- [16] V. Pulkki, Compensating displacement of amplitude-panned virtual sources, in: *Audio Engineering Society 22th Int. Conf. on Virtual, Synthetic and Entertainment Audio*, Espoo, Finland, 2002, pp. 186–195.
- [17] <http://im-research.com/products/designer/>.
- [18] A.J. Berkhout, D. de Vries, P. Vogel, Acoustic control by wave field synthesis, *The Journal of the Acoustical Society of America* 93 (5) (1993) 2764–2778.
- [19] D. Rantzau, U. Lang, R. Ruehle, Collaborative and interactive visualization in a distributed high performance software environment, in: *Proceedings of the International Workshop on High Performance Computing for Graphics and Visualization*, Swansea, Wales, 1996, and <http://www.hlrs.de/organization/vis/covise/features/opencover/>.
- [20] R. Kooima, *Electro*. <http://www.evl.uic.edu/rlk/electro/>.
- [21] http://ivl.calit2.net/wiki/index.php/COVISE_and_OpenCOVER_support.
- [22] <http://www.rcsb.org/pdb/>.
- [23] <http://pymol.sourceforge.net>.
- [24] <http://www.ossim.org/OSSIM/ossimPlanet.html>.
- [25] Aditi Majumder, Rick Stevens, Color non-uniformity in projection based displays: Analysis and solutions, *IEEE Transactions on Visualization and Computer Graphics* 10 (2) (2004).
- [26] M. Agrawala, A.C. Beers, I. McDowall, B. Froehlich, M. Bolas, P. Hanrahan, The two-user Responsive Workbench: Support for collaboration through individual views of a shared space, in: *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH, 1997*, pp. 327–332.
- [27] J.S. Lipscomb, W.L. Wooten, Reducing crosstalk between stereoscopic views, in: *Proceedings of SPIE*, in: *Stereoscopic Displays and Virtual Reality Systems II*, vol. 2409, Feb. 1995, pp. 31–40.
- [28] S. Klimenko, P. Frolov, L. Nikitina, I. Nikitin, Crosstalk reduction in passive stereo-projection systems, in: *Proceedings of Eurographics, 1993*.
- [29] M.S. Puckette, Pure data, in: *Proceedings of the International Computer Music Conference*, International Computer Music Association, 1997, pp. 224–227.



Thomas A. DeFanti, Ph.D., at the University of California, San Diego, is a research scientist at the California Institute for Telecommunications and Information Technology (Calit2). At the University of Illinois at Chicago, DeFanti is director of the Electronic Visualization Laboratory (EVL), and a distinguished professor emeritus in the department of Computer Science. He has researched computer graphics since the early 1970s. His credits include: use of EVL hardware and software for the computer animation produced for the 1977 "Star Wars" movie; contributor and co-editor of the 1987 NSF-sponsored report "Visualization in Scientific Computing;" recipient of the 1988 ACM Outstanding Contribution Award; he became an ACM Fellow in 1994.



Gregory Dawe, MFA, is a Principal Development Engineer at the California Institute for Telecommunications and Information Technology (Calit2). Prior, at the University of Illinois at Chicago, Dawe was Manager of System Services at the Electronic Visualization Laboratory (EVL). He holds a B.A in Design from the University of Illinois at Chicago and a MFA from the School of the Art Institute, Chicago.



Daniel J. Sandin is director emeritus of the Electronic Visualization Lab (EVL) and a professor emeritus in the School of Art and Design at the University of Illinois at Chicago (UIC). Currently Sandin is a researcher at EVL at UIC and at CALIT2 part of the University of California at San Diego. Sandin's latest VR display system is Varrier, a large scale, very high resolution head tracked barrier-strip autostereoscopic display system that produces a VR immersive experience without requiring the user to wear any glasses. In its largest form it is a semi-cylindrical array of 60 LCD panels.



Jurgen P. Schulze, Ph.D., is a Project Scientist at the California Institute for Telecommunications and Information Technology in San Diego, California. His research interests include scientific visualization in virtual environments, human-computer interaction, real-time volume rendering, and graphics algorithms on programmable graphics hardware. He holds an MS from the University of Massachusetts and a Ph.D. from the University of Stuttgart, Germany.



Peter Otto has served on the faculties of Cal Arts and SUNY Buffalo, and as Director of Music Technology at Istituto Tempo Reale, Florence. He is currently on the faculty of UC San Diego's Department of Music as Director of Music Technology where he played a key role in the development of UCSD's Interdisciplinary Computing and the Arts Major (ICAM). In January 2008 Otto was named Director of Sonic Arts Research and Development at UCSD's Calit2. Peter Otto holds a Master's Degree in composition from the California Institute of the Arts.



Javier Girado, Ph.D., is a Staff Engineer at Qualcomm (Graphics team). He earned his M.S degree from the Buenos Aires Institute of Technology (ITBA), held a research fellowship at the Industrial Technology National Institute (INTI), Argentina. He taught at the National Technology University (UTN) and the ITBA. He completed his Ph.D. from the University of Illinois at Chicago (UIC) in 2004. He worked as a Postdoctoral Researcher at the California Institute for Telecommunications and Information Technology (Calit2), University of San Diego, California until 2007. His research interests include virtual reality (VR), auto-stereoscopic displays, computer vision, and neural networks. He specializes in camera-based face detection and recognition to support real-time tracking systems for VR environments, and video conferencing over high-speed networks.



Falko Kuester, is the Calit2 Professor for Visualization and Virtual Reality and an Associate Professor in the Department of Structural Engineering at the Jacobs School of Engineering at UCSD. He received his MS degree in Mechanical Engineering in 1994 and MS degree in Computer Science and Engineering in 1995 from the University of Michigan, Ann Arbor, and the Ph.D. from the University of California, Davis, in 2001. His research is aimed at creating intuitive, collaborative digital workspaces, providing engineers and scientists with a means to intuitively explore and analyze complex, higher-dimensional data. In support of this research, he is developing new methods for the acquisition, compression, streaming, synchronization and visualization of data, including the ultra-high resolution HlPerWall and HlPerSpace visualization environments.



Larry Smarr is the founding director of the California Institute for Telecommunications and Information Technology and Harry E. Gruber professor in the Jacobs School's Department of Computer Science and Engineering at UCSD. Smarr received his Ph.D. from the University of Texas at Austin and conducted observational, theoretical, and computational-based astrophysical sciences research for fifteen years before becoming the founding director of the National Center for Supercomputing Applications (1985) and the National Computational Science Alliance (1997). He is a member of the National Academy of Engineering and a Fellow of the American Physical Society and the American Academy of Arts and Sciences. Smarr is Principal Investigator on the NSF OptiPuter LambdaGrid project and is Co-PI on the NSF LOOKING ocean observatory prototype.



Ramesh Rao received his MS degree in 1982 and Ph.D. degree in 1984, both at the University of Maryland, College Park. Since then he has been on the faculty of the department of Electrical and Computer Engineering at the University of California, San Diego, where he is currently Professor and Director of the San Diego Division of the California Institute of Telecommunications and Information Technology and Qualcomm Endowed Chair in Telecommunications and Information Technology. His research interests include architectures, protocols and performance analysis of computer and communication networks. He was the Editor of the IEEE Transactions on Communications and was a member of the Editorial Board of the ACM/Baltzer Wireless Network Journal as well as IEEE Network magazine. He has been twice elected to serve on the Information Theory Society Board of Governors ('97 to '99 and '00 to '02).