Future Generation Computer Systems 25 (2009) 169–178

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.

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 CAVE[1] prototype was built in 1991, showed full scale ($3m^3$) in public at SIGGRAPH'92[2] 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 m^2 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), lsmar@ucsd.edu (L. Smarr), rrao@ucsd.edu (R. Rao).

[1] The name CAVE™ was coined by the lead author of this paper for the VR room being built at 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.

0167-739X/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2008.07.015

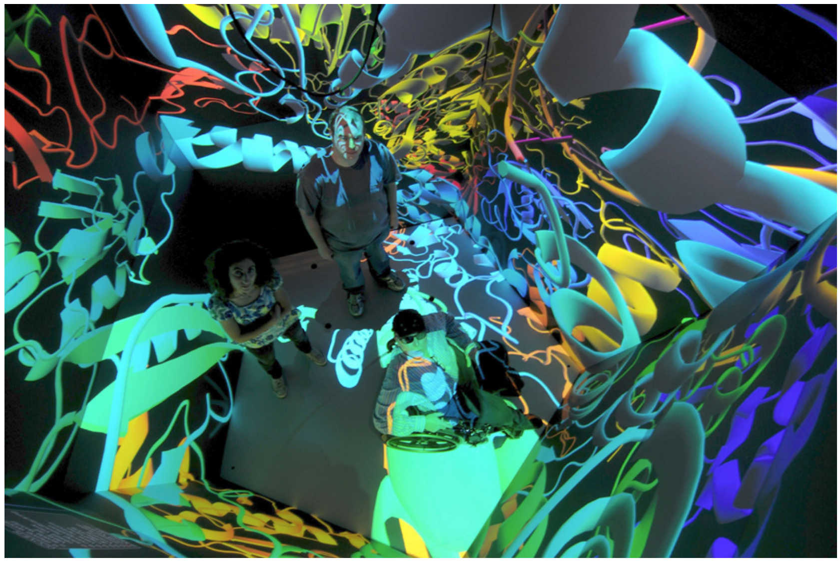

Fig. 1. The StarCAVE from above, looking down on a RNA protein rendering. The still camera taking the picture is not being tracked so the perspective is skewed, but this image shows the floor as well as the walls and shows some of the effects of vignetting and abnormally severe off-axis viewing.

20/140 to 20/200 visual acuity.[3] Besides providingan experience of "low vision",[4] the first CAVEs were relatively dim (the effect was like seeing color in bright moonlight[5]), 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 brighter[7] 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 keystoning 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 built[9] 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 ∼20/40[10] 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 m^2 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.