Michael Cohen*
Spatial Media Group
Computer Arts Lab.
University of Aizu
Aizu-Wakamatsu,
Fukushima-ken 965-8580
Japan

Noor Alamshah Bolhassan
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan
Sarawak, Malaysia

Owen Noel Newton Fernando
Mixed Reality Laboratory
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117574

A Multiuser Multiperspective Stereographic QTVR Browser Complemented by Java3D Visualizer and Emulator

Abstract

To support multiperspective and stereographic image display systems intended for multiuser applications, we have developed two integrated multiuser multiperspective stereographic browsers, respectively featuring IBR-generated egocentric and CG exocentric perspectives. The first one described, "VR_4U_2C" ('virtual reality for you to see'), uses Apple's QuickTime VR technology and the Java programming language together with the support of the QuickTime for Java library. This unique QTVR browser allows coordinated display of multiple views of a scene or object, limited only by the size and number of monitors or projectors assembled around or among users (for panoramas or turnoramas) in various viewing locations. The browser also provides a novel solution to limitations associated with display of QTVR imagery: its multinode feature provides interactive stereographic QTVR (dubbed SQTVR) to display dynamically selected pairs of images exhibiting binocular parallax, the stereoscopic depth percept enhanced by motion parallax from displacement of the viewpoint through space coupled with rotation of the view through a 360° horizontal panorama. This navigable approach to SQTVR allows proper occlusion/disocclusion as the virtual standpoint shifts, as well as natural looming of closer objects compared to more distant ones. We have integrated this stereographic panoramic browsing application in a client/server architecture with a sibling client, named "Just Look at Yourself!" which is built with Java3D and allows realtime visualization of the dollying and viewpoint adjustment as well as juxtaposition and combination of stereographic CG and IBR displays. "Just Look at Yourself!" visualizes and emulates VR_4U_2C, embedding avatars associated with cylinder pairs wrapped around the stereo standpoints texture-mapped with a set of panoramic scenes into a 3D CG model of the same space as that captured by the set of panoramas. The transparency of the 3D CG polygon space and the photorealistic stereographic 360° scenes, as well as the size of the stereo goggles through which the CG space is conceptually viewed and upon which the 360° scenes are texture-mapped, can be adjusted at runtime to understand the relationship of the spaces.

Presence, *Vol. 16, No. 4, August 2007, 414–438*

---

*Correspondence to mcohen@u-aizu.ac.jp.

(a) Cockpit (1st-person) mode (b) Follow (2nd-person) mode

(c) Float (2nd-person) mode (d) Fly (3rd-person) mode

Figure 1. *"Virtua Racing" perspectives (© Sega).*

1 Introduction[1]

1.1 Points of View

A classic example of an exocentric display is a map. But if you allow yourself an imagined out-of-body experience, flying above a landscape to see the world the way it is portrayed in the map, then the map has become an egocentric display. (This is especially easy to accept if the map is replaced by or superimposed upon an aerial photograph of the same area while an avatar of the more omniscient subject remains embedded in the original space.) Such an egocentric display is not categorically distinct from a more exocentric display; rather, it is one extreme of a continuum, along which one can slide back and forth between endpoints that represent egocentric and exocentric impressions or perspectives, as explored by the Worlds-in-Miniature concept (Pausch, Burnette, Brockway, & Weiblen, 1995; Stoakley, Conway, & Pausch, 1995).

Some networked racing simulator arcade games allow each driver to switch between perspective modes. For example, the Sega "Virtua Racing" series has four modes (juxtaposed in Figure 1):

Cockpit (a.k.a. "Drive" or "Dynamic"), shown in Figure 1a, in which the visual presentation is as if the user were inside the car, including the dashboard, steering wheel (with driver's hands), and sometimes rearview mirrors;

Follow (a.k.a. "Basic"), shown in Figure 1b, in which the driver's perspective is yoked just behind and above the vehicle, tracking synchronously;

Float (a.k.a. "Predict"), shown in Figure 1c, in which the virtual camera position is well above the car, still orienting up on the display with forward from the driver's point of view; and

Fly (a.k.a. "Expert"), shown in Figure 1d, in which the monitor tracks the car as if from a blimp, clearly showing one's own car in the context of the field.

---

[1] For extensive online references to all companies, software, and websites associated with this article, please see the online version at http://www.u-aizu.ac.jp/~mcohen/welcome/publications/VR4U2C+JLaY.pdf

Table 1. *Continuum of Navigation Modes*

Perspective		Person	Proximity	Objectification	Virtua Racing	QTVR
Egocentric	Endocentric	1st	Proximal	Reflexive	Cockpit (drive or dynamic)	Panorama
	Tethered or yoked	2nd	Medial	Imperative	Follow (Basic), Float (Predict) Fly (Expert)	Displaced camera distortion: hyperbolic horopter characteristic of unaligned pivot point and focal point
Exocentric		3rd	Distal	Transitive		Turnorama

Even though the simulator's radio buttons select a predetermined degree of immersion, drivers may switch modes during a race, and the visual display slides seamlessly between them, by soaring a virtual camera through the CG (computer graphic) raceway. Further blurring the sampled/synthesized distinction, separate monitors for spectators can show live video of the human drivers, panning shots of the lead car, static shots of strategic curves, and instant replays of crashes (Cohen, 1994, 1998).

We use the word egocentric (centered on the self) to denote displays logically centered on an avatar or position associated with a given user. Such perspectives include both first person and second person metaphors. We reserve the neologism endocentric (centered within) for intimate, strictly first person perspectives, with no explicitly displayed representation of, for example, the user's avatar's head. "Egocentric," then, spans "endocentric" and (for lack of better words) "tethered" or "yoked," and stands in contrast to "exocentric" (centered on the outside), which describes perspectives independent of an avatar. This terminology is summarized by Table 1.

1.2 Image-Based Rendering (IBR)

Interactive computer-generated imagery (CGI, not to be confused with "common gateway interface," as on a web server with soft URLs) can be classified into two main approaches, depending on whether the visual image data is sampled or synthesized. A straightforward methodology, like mesh-generating CAD (computer-aided design), is often inappropriate for geometric modeling of large-scale and complex areas like a cityscape. An IBR (image-based rendering) approach—including techniques such as HoloMedia, Light Field, Lumigraph, Delta Tree, and ray-space methods (Sakagawa, Katayama, Kotake, & Tamura, 2002)—uses 2D information instead of 3D, and the images are photorealistic. The large archive of old media (movies and videos) has great potential for extraction of components for new worlds. These tradeoffs are summarized by Table 2. These technologies can be thought of as endpoints on a spectrum (Hirose, 1997):

algorithmic (polygon) approach
  CG with only 3D models
  CG with texture mapping (of which Just Look at Yourself! is an instance)
  pseudo-3D from 2D part arrangement (of which VR_4U_2C is an instance)
  pseudo-3D from 2D image interpolation
image-based approach

A good analogy for such techniques comes from the domain of synthetic speech: continuous speech samples can be produced either by using physically-based models for speech synthesis, as in a typical vocoder, or by editing and splicing prerecorded samples, as by the vari-

Table 2. *Polygon vs. Image-Based Rendering*

	Polygon	Image-based
Object	Generic: any worlds or objects	Geometry model implicit
Interaction	No limitations	Limited
Quality	Variable	As good as converted 2D media
Limiter	Algorithm-intensive	Data-intensive

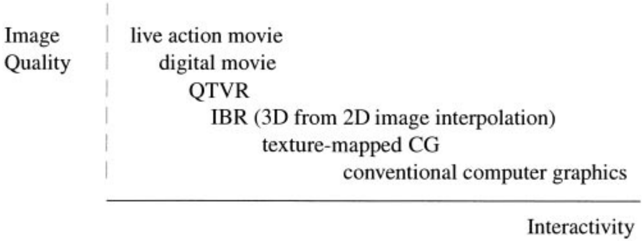


Figure 2. *Image quality vs. interactivity.*

ous synthetic singers based upon the Yamaha Vocaloid. Another way of thinking about these various techniques is to array them according to axes corresponding to image quality and interactivity, as in Figure 2.

1.3 Augmented and Mixed Reality

The basic idea of augmented reality (AR) or mixed reality (MR) in graphics is spatial sensation via 2D information upon 3D models (Milgram & Kishino, 1994; Milgram & Coquhoun, 1999). The virtuality continuum observes that the real/virtual dichotomy is not sharp, but interpolatively smooth. Idealized notions of reality and virtuality can be thought of as endpoints on a continuum, an instance of the former approach corresponding to a see-through display, an instance of the latter to texture-mapped IBR. What we think of as the real world is full of virtual information (from televisions, telephones, computers, etc.), and virtual environments have lots of artifacts of the physical world (like gravity). In the visual domain, techniques associated with AR/MR overlay CGI on top of a real (photographic) scene, or composite sampled data into virtual scenery. Analogously, in the audio domain, computer-synthesized or -manipulated sounds can be mixed on top of a natural ambient soundscape or into directly acquired channels. Sampled data (like that captured by a camera or microphone) is combined with synthesized data, and presented to a human user.

1.4 Panoramas (Cycloramas, or Panos) and Turnoramas (Object Movies, or Turnos)

Apple's QuickTime VR (QTVR) (Chen, 1995) is one of many computer-based interactive image-based technologies, presenting both panoramas (Benosman & Kang, 2001; Jacobs, 2004) and turnoramas (a.k.a. object movies) to support browsing through virtual environments as well as inspection of virtual objects. Internally, the representation of a panorama is different from that of a turnorama. As presented by Table 3, a panorama, normally experienced endocentrically—that is, from a fixed but rotating viewpoint—is stored as a single image, while a turnorama, normally experienced exocentrically—that is, viewing a rotating object from a static viewpoint—is represented as an array of images. A cyclorama or panorama source image is simply panned horizontally, tilted vertically, and zoomed longitudinally, whereas a turnorama array is referenced by an index calculated from perspective state. Highlighting this distinction, a multi-headed display configuration (with several monitors arranged panoramically and displaying separate windows) is natural for panoramic images, but manifests a cubist quality when exhibiting turnoramas, as shown later in Figure 15b. Table 4 reviews the pertinent spatial di-

Table 3. *QTVR Representation vs. Experience*

		Representation	
		Cyclorama or panorama (single image)	Turnorama (array of frames)
Experience	Outward-looking	Ordinary panorama	Panorama as turnorama
	Inward-looking	Turnorama as panorama: unnatural	Ordinary turnorama
Rotation		Subject	Object

Table 4. *Physically Spatial Dimensions: Taxonomy of Positional Degrees of Freedom, Including Cinematographic Gestures**

POSITION					
Static (posture)		**Dynamic (gesture)**			
Location (displacement)	***Scalar***	***Translation camera motion***	***Direction (force)***	***Along axis***	***Perpendicular to plane***
Lateral (transverse)	Abscissa x	Sway track (crab)	left → right	x	Sagittal (median)
Frontal (longitudinal)	Ordinate y	Surge dolly	in, forth (fore): advance (thrust) ↖ out, back (aft): retreat (drag)	y	Frontal (coronal)
Vertical (height)	Altitude z	Heave boom (crane)	up: ascend (lift) ↑ down: descend (weight)	z	Horizontal
Orientation or attitude		***Rotation***		***About axis***	***In plane***
Elevation or altitude	ϕ	Pitch (tumble, flip) tilt	dive/climb	x	Sagittal (median)
Barrel roll	ψ	Roll (flop)	left/right	y	Frontal (coronal)
Azimuth	θ	Yaw (whirl, twist) pan	CCW/CW	z	Horizontal

*Axis descriptions assume right-handed coordinate system with z gravitational up.

mensions and juxtaposes cinematographic correspondences (Arijon, 1991; Vineyard, 2000).

Post-production authoring tools (including VR Worx) can reformat a QTVR movie as either a pano or a turno, so a complete taxonomy should include how the source material was captured, how the authoring production represents it, and how the user experiences it. Ordinary QTVR movies comprise the main diagonal of the central matrix in Table 3. Because of inherent limitations to the panoramic paradigm, precluded by the lower left quadrant, it is generally impossible to present a turno as a pano. (Such an eversion is like the Mapparium at the Christian Science Center in Boston, which presents an inside-out perspective of the Earth that preserves the familiar outside-in features.) However, as a turnorama generalizes the browsing model, it can be used to extend panoramic perspectives—including dynamic effects, non-spatial dimensions, and other interesting experiences (Kitchens, 1998).

For example, a calibrated-rotation turntable and compatible software with digital video camera interface can be used to capture a sequence of stills to make a turnorama (Cohen, 2005a). Normally such capturing and experiences are exocentric, but we put the camera on the turntable to capture an animated panoramic sequence (Cohen, 2002). (Such capture will be less awkward as Bluetooth or other wireless interfaces obviate the cumbersome cable between camera and computer, currently held by an operator who must stride briskly to stay behind the lens.)

Another paradigm-stretching instance (see Cohen, 2005b) is a dental x-ray shot using a CDR (computed dental radiography) panoramic system (the Siemens/Sirona Orthophos 3), which revolves a Röntgen emitter around a patient's head. The resultant 180° volumetric panorama, like that produced by CAT (computer axial tomography), was therefore essentially captured like a turnorama (since the x-rays were pointed at the center of rotation, in the middle of the mouth), but has the single-image form of a panorama. Because of the radial orientation of the outside-in capture, image pairs juxtaposed with this technique are only pseudostereoscopic, contrastable to true stereoscopy as described in the following sections.

1.5 Stereography

Stereography exploits human sensitivity to differences between the parallax views of the world afforded by having two eyes slightly offset from each other. Binocular depth perception results primarily from the processing of the binocular disparity between corresponding points in the two retinal images that result from viewing a scene containing objects at different depths (McAllister, 1993; Davis & Hodges, 1995; Martens, McRuer, Childs, & Virree, 1996; Nagata, 2002). In everyday viewing, this is accomplished by converging the eyes on a target, which brings the fixated target point on the two retinal images into alignment at zero disparity. Objects nearer to the viewer than the plane of fixation exhibit crossed disparity, while objects further from the viewer than the plane of fixation exhibit uncrossed disparity. The human binocular visual system is exquisitely sensitive to small variations in disparity, and can resolve disparities as small as 20° arc in visual angle. (For more detailed information about human binocular visual sensitivity related to stereography, see Diner & Fender, 1993.)

Just as in natural binocular viewing, disparities within stereoscopic image pairs give rise to depth percepts that can appear very natural. Spatially sampled visual data from naturally occurring scenes, presentable via stereo-image pairs for stereoscopic viewing, can be captured simply by using two cameras separated by a suitable interocular distance. (The mean distance between human eyes, nominally 65 mm, or about 2.6 in., is a natural choice for many scenes.) These stereo-image pairs may be readily viewed through an appropriate stereographic image display device (selecting between two images for the respective eyes). Interactive stereographic display of such sampled images is difficult if users are allowed to freely change the viewpoint, although some IBR techniques could potentially be applied to such stereoscopic generation (Katayama, Tanaka, Oshino, & Tamura, 1995; Shade, Gortler, He, & Szeliski, 1998; Rademacher & Bishop, 1998; Oliveira & Bishop, 1999; Snavely, Seitz, & Szeliski, 2006).

In contrast, interactive stereographic 3D-model-based CGI is relatively easy, as the synthesis of a second image is basically free, requiring only a second projection and rendering for a virtual camera viewpoint displaced by a simulated interocular distance. Millions of people have experienced such stereographic 3D-CGI systems, and the recreation marketplace has already been penetrated by interactive computer games that offer stereographic viewing options. Some modern GPUs (graphics processing units, formerly known as co-processing accelerators) even feature native stereo output: the geometry sent from the CPU is elegantly rendered as a stereo pair in hardware.

Interactive stereographic display of sampled stereo-image pairs is more difficult because stereo pairs from each desired viewpoint must be captured. The binocular disparity presented in stereo-image display is a consequence of interocular displacement during capture, and if the multiple viewpoints are sampled with capture pairs that are not properly oriented (i.e., facing in nearly the same direction), then binocular disparities will be distorted. Further, if the stereo camera pair moves closer to an object, binocular disparities will be differentially modulated in comparison to more distant objects.

Of course, when a single camera moves through a 3D scene, motion parallax is also generated. Displacing a camera laterally provides particularly strong information about the depth of visual objects via motion-generated parallax views. Distant objects appear to move relative to nearer objects, traveling in the same direction as the camera. If a single camera dollies in or out of a scene, closer objects will loom large in the camera's field of view, but more distant objects previously occluded may also be revealed. Backing up and zooming in yields a compressed sensation. Such translation-dependent effects are not supported in ordinary applications built using the QuickTime multimedia architecture, but our unique multinode implementation provides these effects as a natural consequence of a technique that enables stereoscopic viewing of panoramic images displayed using an extension of QTVR. The following section describes the context and motivation for the development of this stereographic QTVR (which we have dubbed SQTVR) feature.

1.6 SQTVR: Stereographic QuickTime Virtual Reality

One of our motivating goals has been the idea of giving pictures depth, like that exploited by the fictional Esper from the movie *Blade Runner* (Dick & Scott, 1982), which could extract almost limitless information from a single hyper-still, allowing users to look around corners and behind walls, seeing previously occluded objects. Such photographic omniscience recalls the successive magnifications used in the movie *Blowup* (Cortázar & Antonioni, 1966). Deep panos captures the idea of aligned multinode movies, which interpolate between the 2D cubical/cylindrical/spherical geometry of QTVR and the 3D geometry of CAD. We have implemented a browser, "$\mathrm{VR}_4\mathrm{U}_2\mathrm{C}$" ('virtual reality for you to see'), that performs track- and dolly-enhanced QTVR, with the additional feature that panoramas can be viewed stereographically, along with a sibling browser "Just Look at Yourself!" that uses the same panoramic images as texture maps to visualize and emulate the SQTVR projection. For stereo panoramic browsing such as that provided by SQTVR, many viewpoints must be captured, as the vector between the stereo eyepoints must always be almost perpendicular to the view direction so that appropriate disparity between the respective sides of the stereo pair is maintained. For degenerate counterexample, if the viewing direction lies along the edge between two eyepoints, no binocular parallax is possible.

2 Related Research

The multinode SQTVR technique can be compared and contrasted to other extended IBR and panoramic methods. Our straightforward approach recalls the classic Movie Map (Lippman, 1980; Naimark, 1997). Like the Sea of Images approach (Aliaga, Funkhouser, Yanovsky, & Carlbom, 2003; Aliaga, Yanovsky, Funkhouser, & Carlbom, 2003), our technique relies upon dense sampling, constructing what can be thought of as a 4D approximation to the generalized plenoptic function, describing radiance arriving at any point (x, y) in a plane from any direction (θ, ϕ). However, our approach differs in that we invoke no view interpolation (Chen & Williams, 1993) or image warping (morphing; Seitz & Dyer, 1996), as performed by several browsers, of which McMillan and Bishop (1995a) is an early representative.

Tightening the alignment between the sampled and synthesized scenes is also beyond our current interests. A more seamless integration would require reducing registration errors, both static and dynamic (which implies low-latency tracking), as well as rendering errors.

Many modern first-person shooter (FPS) games recreate an actual place (such as Activision's True Crime: New York City). In contemporary photogrammetric practice, a standard approach is to capture enough panoramas in a limited area such that all surfaces are unoccluded, then mark up a geometric model to reference those panos to support continuous walkthroughs. The tradeoff is that a photogrammetric process requires a skilled modeler and provides continuous viewing, whereas our SQTVR capture process is simpler but more tedious, and yields pure stereographic pairs.

Rather than making a mosaic of captured video (Bartoli, Dalal, & Horaud, 2004) or accumulating a vertical pixel column from successive frames of a continuous video in an arbitrarily long route panorama (Zheng, 2003), our panoramas are discretely captured and compiled into multinode movies. We use a set of panoramic images, rather than an array of cameras, as in Matusik and Pfister (2004). The mirror-based or coincident projection techniques we use require only a single image for each channel, and therefore we require no conformal texture-mapping, like that used by Raskar et al. (2003); Raskar, van Baar, Willwacher, and Rao (2004), or edge-blending techniques, like that used by Hayashi, Nakaizumi, Yano, and Iwata (2005). In these senses, ours is a brute force approach.

One of our explicit goals was the generation of stereographic views (which presumably could be performed by most of these other IBR techniques). Since our SQTVR browser continuously selects two such normal panos from such a self-contained bundle at runtime, a pair is dynamically assembled rather than intrinsically stereographic. This level of indirection, controlling selection of panoramas, allows tracking and dollying and is the key to our IBR system, especially as the control can be networked and distributed to other complementary interfaces.

2.1 Panoramic and Turnoramic Displays

Many products and applications have been developed for binocular and panoramic capture, stitching, and display, including Cave-like (Cruz-Neira, Sandin, & DeFanti, 1993) spatially immersive display, and artistic installations (Mohr, 2004). Advanced and expensive equipment, such as Panoscan or FullView (Nalwa, 2000a, 2000b) cameras, can be used to get very high-resolution panoramic images, and more affordable solutions have also been commercialized, such as the Eizoh Hyperomni Mirror system. Users may also prefer to use either a slit-scan panorama camera or coupled multiple cameras (Tzavidas & Katsaggelos, 2005). Currently commercially available display systems include the Elumens VisionDome and VisionStation, the Matsushita Panasonic Shiodome CyberDome and CyberDome 1800, and the SpinDome. Toshiba is developing a personal (6 lb) head-tracked viewing dome helmet with a 16 in. screen. Such immersive displays envelope a user or a group of people without restrictive head-mounted displays. A screen is arranged (often hemispherically) to fill the field of view of the participants, creating a sense of immersion in the same way that large-screen cinema draws an audience into a scene. An observer loses most of the cues regarding display surface position, such as screen edges, and perceives 3D objects off the surface of the screen. A dome allows freedom of head motion, so that the observer can change direction of view and yet still have vision fully encompassed by the image. Such display systems can enhance visualization power in the fields of space planning, manufacturing and design, simulation and training, entertainment, education, and medical research.

One of many advanced systems for volumetric display (Favalora, 2005) is the Perspecta Spatial 3D, a 20-in. dome displaying full-color and full-motion interactive images that occupy a volume in space, giving users an all-encompassing view without goggles. Other revolutional systems that enable users to view an object from any angle are the SeeLINDER (Yendo, Kawakami, & Tachi, 2005), a cylindrical 3D color television, and the Spinning Display (Maeda, Tanikawa, Yamashita, Hirota, & Hirose, 2004), a flat liquid-crystal display (LCD) revolved mechanically by a stepping motor. These kinds of display systems can be used in many fields, including medical imaging, geophysical research, defense/security, and biotech.

In contrast to such display systems that are very costly and normally used for specific purposes—such as large theater halls, automotive and aerospace designs and presentations, special training, and so on—the multi-window and multimonitor system described here offers panoramic and turnoramic browsing features that can be enjoyed by users of normal computers, especially if they have an internet connection and/or a multi-display system.

2.2 Stereo Panoramas

2.2.1 Capture. There have been many techniques proposed to prepare stereo panoramas. A traditional technique (Huang & Hung, 1998) uses a rotating stereo head with two displaced cameras, stitching the captured multiple rectangular images separately into two panoramic mosaics. However, this technique demands extra time and effort to prepare a good stereo pano, requiring several rectangular images for each side of the stereo pair. As severe problems of parallax and scale changes are difficult to avoid, another technique for generating stereo panoramas combines multiple stereo camera units, like that employed by the Stereo Omnidirectional System (SOS) (Shimada, Tanahashi, Kato, & Yamamoto, 2001), which has sixty color CMOS cameras and omnidirectionally captures raw and depth images synchronized in real time. Of course such an array is prohibitively expensive for general users.

Alternatively, a more convenient technique (Naemura, Kaneko, & Harashima, 1998; Peleg & Ben-Ezra, 1999; Kawakita, Hamaguchi, Tsukamoto, & Miyazaki, 2000) uses video cameras to prepare a stereo panorama by making a mosaic of interlaced vertical narrow (left and right) strips on every video frame separately. It also enables control of stereo disparity, giving larger baselines for faraway scenes and smaller baselines for closer scenes. A related technique, Omnistereo, proposes a special camera catadioptric (reflection and refraction via mirror and lens) assembly, including a spiral lens or mirror (Shum, Kalai, & Seitz, 1999; Pritch, Ben-Ezra, & Peleg, 2000; Peleg, Ben-Ezra, & Pritch, 2001) or a catacaustic of a cylindrical (or cylindroidal) mirror (Tanaka, Hayashi, Endo, & Tachi, 2003). The BeNoGo project also developed novel camera technologies to allow nearly photorealistic navigation.

The technique described here explores an alternative method, using multiple images captured by an omnidirectional camera (Nayar, 1997), allowing users not only to view and experience stereo panoramic scenes but also to enjoy stereoscopic navigation. Our technique does not require the stitching together of multiple captured views, so is quite distinct from “panoramic image mosaics and environment maps” techniques for IBR systems, such as those proposed in Szeliski and Shum (1997).

2.2.2 Navigation/Walkthroughs. Current image-based panoramic visualization browsers provide a way to experience virtual and real-world environments using cubical, cylindrical, or spherical panoramas. As viewpoint is positioned at the center of a panorama, a user can pan to the left and right, tilt up and down, and zoom in and out. Since the viewpoint is fixed, a navigation/walkthrough experience (Kotake et al., 2001) can be achieved by constraining the camera movement to only particular locations where the panoramic images were captured, as briefly described in Chen (1995), which suggested that mobile viewpoints in space could be snapped to a nearest grid point to approximate motion.

Image morphing techniques can be used to allow interactive panoramic navigation or walkthroughs of virtual and real-world environments, including general morphing (Hirose, Watanabe, & Endo, 1998), mesh triangulation and morphing (Darsa, Silva, & Varshney, 1997), epipolar geometry (Chiang, Way, Shieh, & Shen, 1998), and pixel-based (McMillan & Bishop, 1995b) and triangle-based (Fu, Wong, & Heng, 1998; Fu & Heng, 1998; Fu, Wong, & Heng, 1999; Chan, Fok, Fu, Heng, & Wong, 1999) image warping techniques, which cause a source image to melt, dissolve, and rearrange itself to form a target. However, these techniques require dense depth or correspondence information between images to perform morphing and/or warping during walkthrough, which association is tedious and sometimes impractical for real-world photos.

Mesh-based techniques assume that corresponding meshes are defined on reference frames, with appropriately matching topologies and control points. The challenge is to warp the perspectives without introducing artifacts, preserving parallel and converging lines. Images are interpolated by converging corresponding control points, with a final step of blending the respective distorted images, perhaps by simply compositing weighted frames. Such techniques work best when the control points are explicitly specified by a human operator, and so are less well suited for the approach described here, which uses many nodes to support stereographic dollying and tracking.

2.2.3 Anaglyphic Stereographic QTVR. Stereographic viewers for panoramic and turnoramic images were invented more than a century ago (Drouin, 1995; Waldsmith, 2002), and computer-presented stereo panoramas and turnoramas have been available for more than a decade. Most of these panoramic stereo displays were available only in anaglyphic format, the binocular effect viewable through eyewear with a red (or amber) filter over one side and a blue (or green or cyan) filter over the other. To prepare such a stereo QTVR panorama or turnorama, pairs of stereo images can be tinted and composited to produce anaglyphic images and then converted into QTVR movies. However, image colors in anaglyphic format do not always properly filter, resulting in leakage and cross-talked ghosts. With the richer channel-respecting viewing techniques that browsers like ours offer, users can view stereoscopically in native full color. A recent technique, related to anaglyphic techniques but featuring more tightly resolved wavelength multiplex imaging (Jorke & Fritz, 2003) driven through binocular eyewear coated with dielectric material sharply multituned to the multifrequency sources, claims fuller-color channels with negligible cross talk.

3 Multimodal Groupware Architecture

As infrastructure for the integration of heterogeneous interfaces, including the multiperspective IBR and CG clients that are the focus of this article, we had developed an architecture and framework (Kanno, Cohen, Nagashima, & Hoshino, 2001; Fernando, Adachi, Duminduwardena, Kawaguchi, & Cohen, 2006) to support collaborative virtual environments (CVEs), allowing distributed users to share multimodal virtual spaces. Our CVE architecture is based upon a simple client/server model, and its main transaction shares the state of virtual objects and users (avatars) by replicated-unicast of position parameters (translation, rotation, plus zoom) to client peers in a session. There is no server caching of state, and changed parameters are immediately redistributed, with the exception of "servents" (server/client hybrids) that act as intermediating proxies for mobile phone clients. The main features of our CVE suite are multimodal signaling, platform independence, and easy network connectivity, as components are built with Java, including JSE (Java Standard Edition: core/desktop; Lea, 2000; Horstmann & Cornell, 2001), Java Sound, JMF (Java Media Framework; Gordon & Talley, 1999) and its media API rival QuickTime for Java (Maremaa & Stewart, 2005; Adamson, 2005), Java3D (Sowizral, Rushforth, & Deering, 1998, 2000; Palmer, 2001; Tanaka et al., 2002; Walsh & Gehringer, 2002; Selman, 2002; Ota, 2003; Hirouchi, 2004), JME (Java Micro Edition: mobile; ASCII Editing Group, 2001; Feng & Zhu, 2001; Yamazaki, 2001; Mahmoud, 2002; Topley, 2002; Vacca, 2002; Knudsen, 2003; Kontio, 2003; Fukuoka, 2004; Li & Knudsen, 2005), and Swing (Walrath & Campione, 1999).

The CVE suite integrated by this framework includes

VR$_4$U$_2$C (Bolhassan, Cohen, & Martens, 2004),
Just Look at Yourself! (Bolhassan et al., 2002),
the Soundscape-Stabilized (Swivel-Seat) Spiral-Spring GUI (Cohen & Sasa, 2000) for modeling azimuthal spatial sound displayed through a rotary motion platform,
the RSS-10 CSD Speaker Array Driver (Cohen et al., 2002; Sasaki & Cohen, 2004) enabling an eight-channel display,
the $^{S}_{c}\mathrm{ha}_{i}\mathrm{r}^{e}$ (for shared chair) Internet Chair (Koizumi, Cohen, & Aoki, 2000; Cohen, 2003; Duminduwardena & Cohen, 2004; Kanno, Fernando, Bolhassan, Narita, & Cohen, 2006) rotary motion platform (shown in Figure 3),
2.5D Dynamic Maps (Mikuriya, Shimizu, & Cohen, 2001),

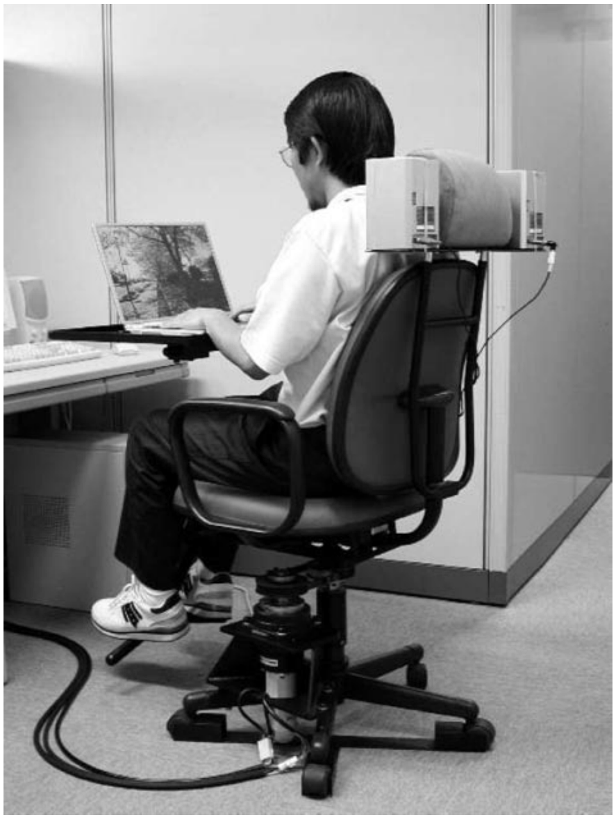


Figure 3. *$^{S}_{c}$ha$_{i}$r^{e} Internet Chair integrated with VR$_4$U$_2$C: The swivel chair is a rotary motion platform. The panorama displayed by the laptop computer (with wireless internet connectivity) is propioceptively anchored, panning opposite to the chair rotation driven by the servomotor and the soundscape displayed through "nearphones," loudspeakers straddling the headrest for binaural display without crosstalk.* (Developed by the first author and Uresh Duminduwardena with the collaboration of Yamagata University and Mechtec.)

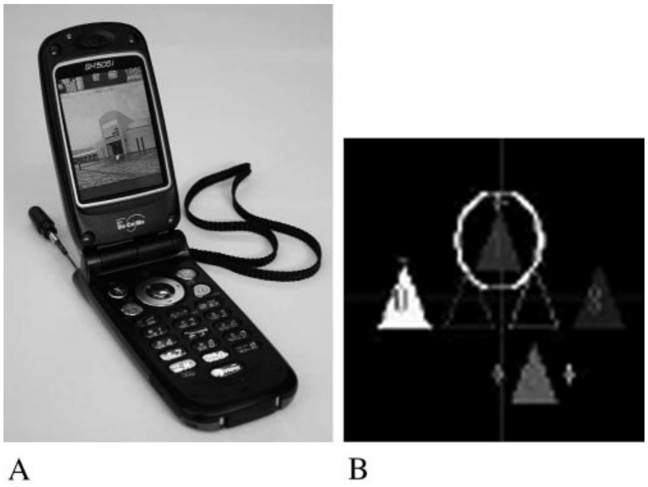


Figure 4. *These two mobile applications for NTT DoCoMo iαppli mobile phones are integrated with "VR$_4$U$_2$C" and "Just Look at Yourself!," and can reflect and control the workstation displays. (a) Egocentric Panoramic Browser μ "VR$_4$U$_2$C": Some mobile phone displays feature autostereoscopic (lenticular) displays. Our dataset and algorithms have been ported to such mobile devices, enabling stereographic viewing of QTVR-like panos and turnos (with quasi realtime network synchronization). (Developed with Etsuko Nemoto.) (b) Exocentric Dynamic Map "$\iota \cdot$ Con": A 2.5D map displays and controls iconic location and orientation, along with narrowcasting attributes. (Developed with Yutaka Nagashima, Makoto Kawaguchi, Gō Saito, and Yoshi Tanno.)*

- the Pioneer Sound Field Controller (PSFC) proxy (Amano et al., 1998) enabling a 15-channel speaker array,
- Java3D widgets for games (Adachi, Cohen, Duminduwardena, & Kanno, 2004) and simulators (Adachi, Iwai, Yamada, & Cohen, 2005; Nagai, Cohen, Moriguchi, & Murakami, 2007), and
- mobile applications, including these MIDlets (JME mobile information device applets) for mobile phones
  - μVR$_4$U$_2$C shown in Figure 4a, and
  - $\iota \cdot$ Con (Kanno, Cohen, Nagashima, & Hoshino, 2001; Ishikawa, Saito, & Cohen, 2005; Fernando, Saito, Duminduwardena, Tanno, & Cohen, 2005; Fernando, Cohen, Duminduwardena, & Kawaguchi, 2006) narrowcasting interface, shown in Figure 4b,

which together echo the egocentric/exocentric differences highlighted by the first two clients, as described following.

The juxtaposition of these stereographic browsers is the focus of this article. "VR$_4$U$_2$C" is a photorealistic, endocentric (1st person), image-based browser allowing panoramic panning, as well as tilting and zooming. "Just Look at Yourself!" is a computer graphic browser, allowing more flexible camera positions, including endocentric (1st person: from the point of view of the avatar), tethered (2nd person: attached to but separate from the avatar), and exocentric (3rd person: totally detached from the avatar) per-

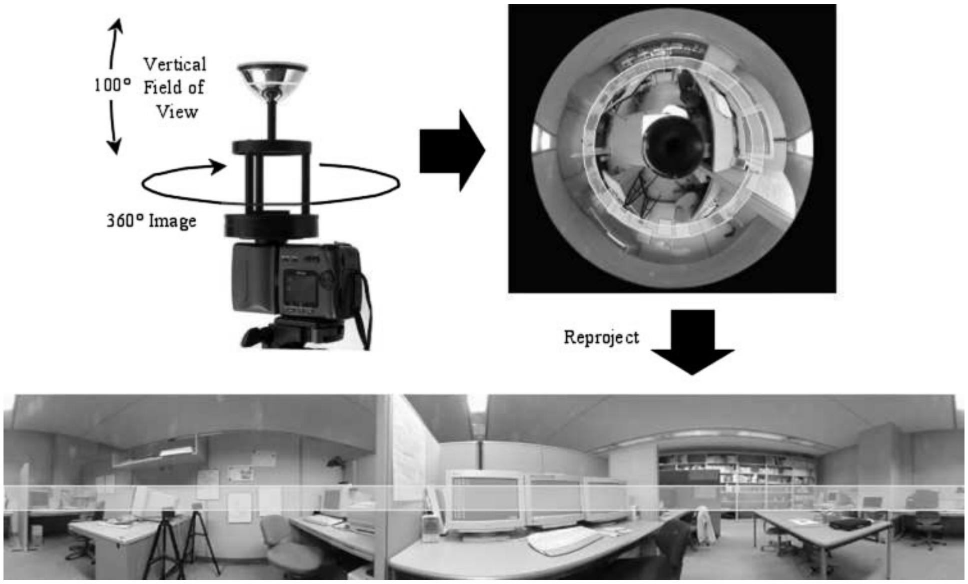


Figure 5. *Preparing a panoramic scene. (Equatorial bands are superimposed to show corresponding horizons in the various projections.)*

spectives (Barrilleaux, 2001, pp. 125–147). They are respectively described in the following sections. Interested readers are encouraged to visit the project web page (see Fernando, 2007) to stream videos of the applications in action, download QTVR panos captured with the processes described here, and install and run the Java applications (server and clients).

4 $\text{VR}_4\text{U}_2\text{C}$: Networked Multimonitor SQTVR with Dollying and Tracking

4.1 Stereographic Panorama and Turnorama Imaging

4.1.1 Monocular Panoramic Capture. Panoramic scenes can be photographically captured using a variety of techniques. We used a NikonCoolPix 990 or CanonKiss Digital SLR digital camera with remote shutter release, Kaidan and EyeSee360 360 One VR optical system (with either CoolPix 990 or SLR bracket mounting kit), and a monopod with bubble level. With a single camera shot, the equiangular mirrored optical system can capture a 360° panorama with a 100° vertical field-of-view (50° above and below the horizon). The camera and lens system allows a very large depth-of-field, as the hyperbolic mirror is already displaced from the actual camera lens, so that even objects close to the standpoint can be in focus. Bundled PhotoWarp software can be used to process the captured “donut” panoramic image (as shown in Figure 5), yielding a QTVR movie or cylindrical image.

4.1.2 Stereo Panoramas: Capture and Dynamic Node Selection. The equipment we use for capturing and preparing stereo panoramic scenes includes all that for ordinary panoramic capture plus a compass and a “capture lattice” as shown in Figure 6. First, a camera atop a monopod at the center of the lattice (`point 1`) is used to take a picture, capturing a complete panoramic scene. As for ordinary panoramic capture, this image is processed and saved as a QTVR movie or cylindrical image, used as the left-eye pano in the simple scenario traced below.

To ensure that the vector between the left-right pairs is perpendicular to the view angle, our system

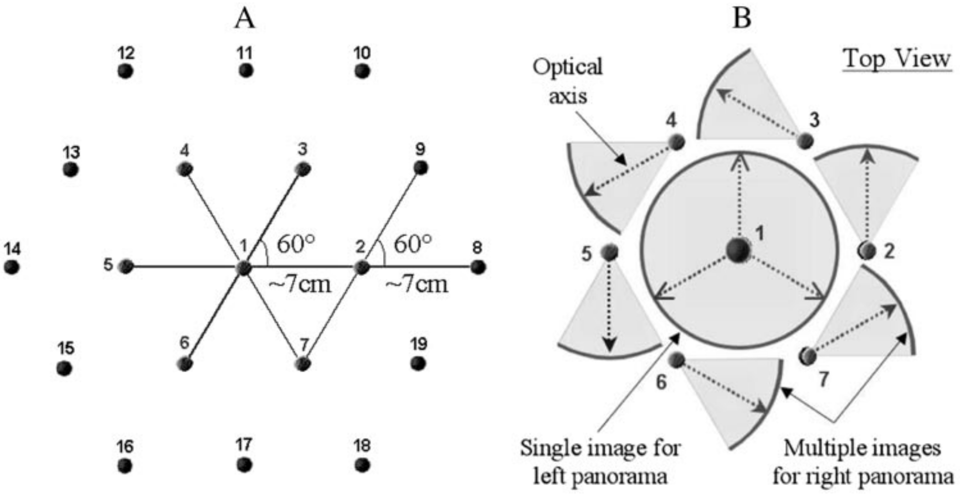


Figure 6. *Stereo panoramic capture and node selection: capturing discrete panos at vertices of plane-tessellating equilateral triangles with edges equal to the interocular distance enables dynamically selected stereo pairs. (a) Capture lattice. (b) Node selection.*

establishes a pivot point, about which one side of the left-right pair revolves. (Such a concept is familiar to sports, especially regarding placement of the feet, as when a baseball fielder reaches for a throw on a force out, an Ultimate flying disc player stretches out for a throw, or a basketball player avoids traveling.)

Therefore, to prepare the right-eyed side of a stereo panoramic scene, another picture is taken after displacing the monopod an interocular distance (65 mm, to `point 2`), adjusting so that the compass attached to the monopod indicates the same direction as the capture for the left-eyed panorama, ensuring that every captured pano will align with the original. This process is repeated until all necessary nodes have been captured, and these panos are compiled into a multinode QTVR movie. Such node arrangement represents a partial tiling of the horizontal plane with equilateral triangles, the simplest regular tessellation that allows stereographic node pair selection. Happily the resultant 60° sectors allow smooth azimuthal transitions.

The numbered nodes in Figure 6b index the rendering of a multinode SQTVR movie, allowing dollying and tracking for stereographic standpoint navigation using a variety of steering techniques. To render a stereographic scene with an initial pan angle (0°), VR_4U_2C displays `node 1` as the left-eye pano and `node 2` as the right-eye pano, as in Figure 7. While panning to the left or right, VR_4U_2C monitors the azimuth and changes the right-eyed panorama to the appropriate node as the angle reaches a certain threshold. If the user dollies to the right, the left-eye panorama will change to `node 2` and the right-eye panorama to `node 8`; if the user dollies to the right-front, the left-eye panorama will change to `node 3` and the right-eye panorama to `node 9`; and so on. Figure 8 illustrates how left and right stereo paths can be arranged straddling a free planar path.

Besides stand-alone control of each viewer, our distributed control system allows locally generated (from sibling clients on the same computer) or network events (from distributed peers) to steer through multinode panoramas. A simplex (half-duplex) mode enables coupling between local control and display, while a (full) duplex mode disables such immediacy, relying instead on returned network events to update the browser. This scheme accommodates network delays and client latency, synchronizing multimodal output. For particular instance, the ${}^{S}_{c}\mathrm{ha}_{\mathrm{i}}\mathrm{r}^{\mathrm{e}}$ Internet Chair has significant sluggishness, a consequence of me-

Figure 7. *Example of left/right stereo images. Note binocular parallax viewpoint displacement, causing offset of objects like the paper on the left edge.*

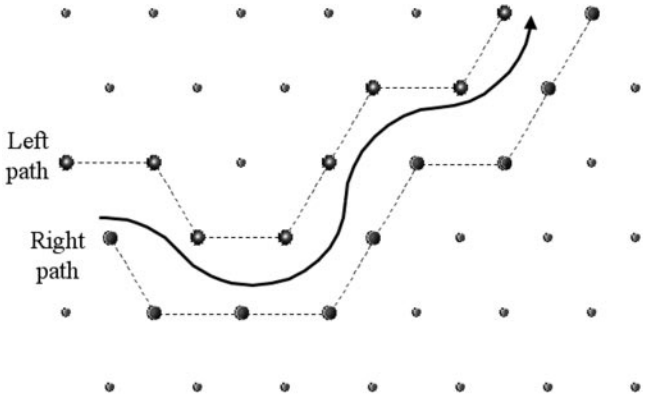


Figure 8. *Left/right stereo paths determined from walkthrough meander.*

chanical inertia and user comfort. When its software controller receives a target azimuth, it sends continuous updates on a separate channel while twisting the motion platform toward the goal, coupling the visual, auditory, and proprioceptive displays.

Revolution about a point incurs changes in both orientation and location. Unlike an ordinary panorama that only rotates on a single point changing orientation (yaw), the multinode stereographic feature allows objects to loom, relative angle subtense changing for dolly but not for zoom. Our implementation has a “winking” artifact, caused by asynchronous updates of the lateral panes, and the stereo stitching is not entirely seamless; there are subtle but distinct hiccups when switching between the multiple panorama nodes used for one side of a stereo pair. These discontinuities are like the judder (a portmanteau word combining jitter and shudder) observed during slow-motion video playback (Watkinson, 2001, pp. 382–383). Such effects are typically not noticed by unsophisticated users until they have been pointed out, but are usually discernible afterwards. A higher spatial sampling density (more capture-intensive) would alleviate such side effects. Alternatively, rotation about the center point of the interocular axis, the Cyclopean eye, could provide smoother rotation with less obvious sway and surge.

The quantized node lattice circumference, built on a triangular grid, grows linearly with radius r, and the hexagonal area A_r, the number of nodes in the interior, grows quadratically:

$$A_r = 1 + \sum_{n=1}^{r} 6n = 1 + 3r + 3r^2$$

We captured a 100-node lattice in our lab, enabling 5-step planar dollying or tracking in whatever horizontal

Figure 9. *Equipment for capturing turnoramas.*

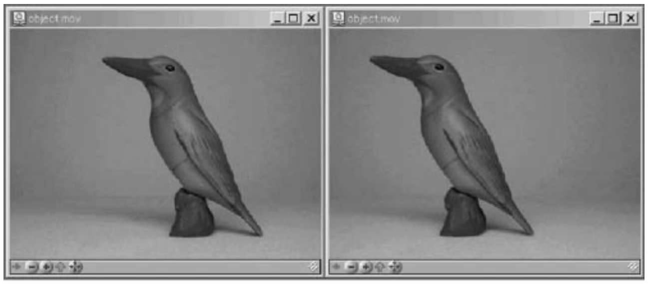
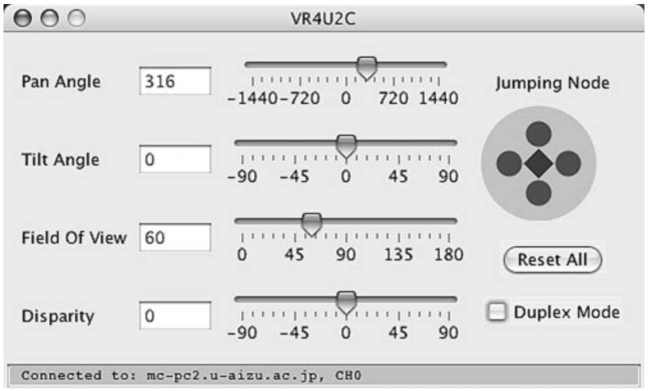


Figure 10. *Example of left/right stereo turnoramic images.*



Pan Angle 316

Tilt Angle 0

Field Of View 60

Disparity 0

☐ Duplex Mode

Connected to: mc-pc2.u-aizu.ac.jp, CH0

Figure 11. *VR$_4$U$_2$C navigator with continuous and numeric controllers including dolly and track rosette.*

directions, allowing users to move around and view a scene from multiple stereographic standpoints.

The QTVR headers are modest, and there is no internode compression, so the size of a multinode file is basically the sum of its constituent panos. The 100-node pano described above is about 28 MB—reasonable for a workstation, but heavy for contemporary mobile phones.

4.1.3 Stereo Turnoramas: Capture and Disparity Adjustment. As mentioned earlier in Section 1.4, turnoramas (object movies) can be photographed using, for example, a FireWire-enabled digital video camera, a modern computer, and 3D Object Imaging Kit (a 3D object capture solution that includes a calibrated-rotation turntable and the Autolycus SpinImage DV software). An object of interest is centered on the turntable, and a video stream of the spinning object is captured by the camera, as shown in Figure 9. Software can convert the video stream into a QTVR turnorama or a set of still images. Multinode bundles such as those needed for panos are not needed, as the respective viewing angles of the two eyes reference interoccularly offset object views in the azimuthal image vector for pseudostereopsis.

Users can browse QTVR turnos stereoscopically (as shown in Figure 10) through VR$_4$U$_2$C's multiple synchronized window feature, presenting a stereo pair with phase adjusted as disparity, typically 5° or more. Animated presentation of stereoscopic animated turnoramas is most effective for objects that are spun horizontally and lightened equally, preserving photometric alignment.

Our browser can be used to produce a pseudostereo pair by simply offsetting otherwise identical views of such single-frame images, like the dental x-rays described earlier. VR$_4$U$_2$C allows adjustment of turnorama binocular disparity, as shown in Figure 11, from normal through super- to hyperdisparity, to expand and contract the depth modulation as in the Disneyland attraction "Honey, I Shrunk the Audience." Such adjustment causes a subjective shrinking of the perceived world, as the user unconsciously recalibrates perception to the natural interocular distance. (However, due to

Figure 12. *Multiple synchronized windows.*

the different nature and representation of panos and turnos, such parameterization applies only to turnoramas.) The disparity slider goes past 0° (which datum is useful for calibrating a stereographic viewer, including angle trims on mirrored stereographic viewers) to negative, hypostereo values.

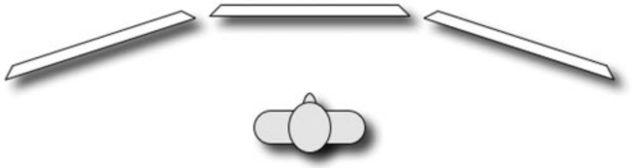
4.2 Client Integration

As one of many integrated clients, VR$_4$U$_2$C connects to a session server, as previously described in Section 3, to synchronously exchange parameters with sibling clients, including any other instances of VR$_4$U$_2$C. If a user changes tilt or pan angles, dollies, tracks, or zooms, the new perspective is multicast through the server. Upon receiving orientation, location, or zoom values from a session server, VR$_4$U$_2$C refreshes its state, assigning pitch to tilt angle value and yaw to pan angle value, and translating the stereo standpoint appropriately. (However, the program will only cache roll values, as our system has no provision for making a displayed image barrel roll.) VR$_4$U$_2$C also gets zoom values and left and right node IDs from the session server, assigning them to the field of view and current lateral pano or turno indices.

4.3 Browser Features

4.3.1 Multiple Synchronized Windows/ Monitors. With the multiple synchronized displays feature, users may use VR$_4$U$_2$C to open and browse QTVR panos and turnos across multiple frames/ windows with the same or different viewing angles. Users can also deploy a multidisplay system for viewing multiple screen-sized windows on separate monitors. If desired, one could view many aspects of pano scenes on multiple monitors, which arrangement might be called a “panoramic panorama.” Users may set a mullion width value for arraying multiple windows contiguously, and model the frames and bezels of monitors as mullion-like borders. (This value is specified in pixel units, so that it can be used for view width, mullion width, and pan angle calculations.) Figures 12 and 13 illustrate how multiple synchronized windows and monitors of a panorama can be arrayed; the horizontal gap between them is the mullion width. This idea is extrapolated by the concept shown in Figure 14. Further, users can arrange multiple monitors cyclically (especially flat-panel LCD monitors with a compact footprint), as shown in Figure 15, to display a turnorama through a complete cycle of viewing angles simultaneously, which might be called a “turnoramic turnorama,” as in a virtual showcase.

4.3.2 Stereographic QTVR Viewing with Navigation. We have generalized the discrete capture technique to allow not only stereographic capability, but also dollying and tracking, true virtual camera motion in IBR (with admittedly very limited perspective latitude). Stereographic (and dolly-able) browsing uses the multiple synchronized windows feature of our system. To view SQTVR movies, users may select one of three modes: `parallel`, `cross-eyed`, or `over/under`. If the `parallel` option is chosen, an opened

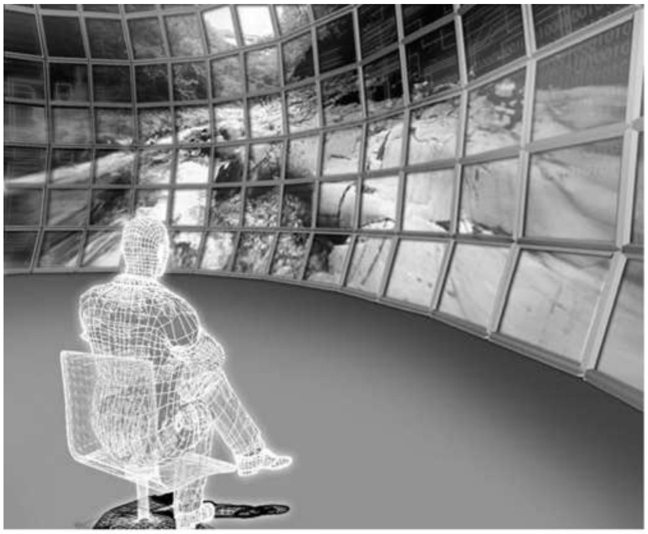


(a) Schematic

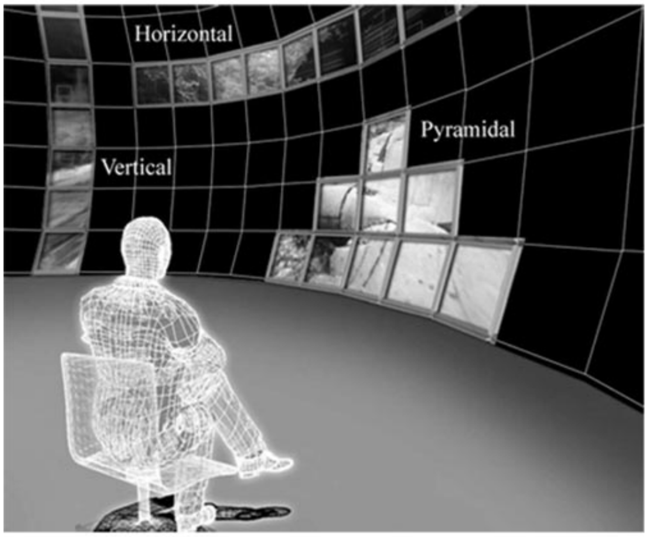
(b) Typical view

Figure 13. *Multiple synchronized monitors: panoramic panorama.*

movie pair is displayed with the view for the left eye on the left, and that for the right eye on the right. This style is the most common, aligning as it does with the natural arrangement of the eyes, and is used in many visualization extensions, including the stereographic feature of LiveGraphics3D, the Mathematica stereographic manipulation applet extension. With the `cross-eyed` option, an opened movie pair is displayed side-by-side as well, but with the stereo pair swapped. For a special viewer called Leavision, the `over/under` option arranges a stereo pair vertically (with the movie for the right eye above that for the left).



(a) Immersive Imagery

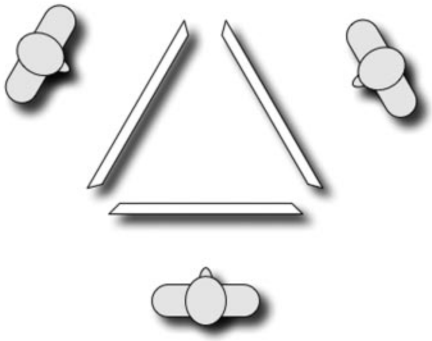


(b) Tiling Styles

Figure 14. *Graphic depiction of a curved "video-wall" configuration of visual display. (Graphics by "Eyes, Japan.")*

A pair can be free-viewed without any viewing aid, but such fusion takes some practice as it involves paralleling or crossing one's eyes slightly. Eye strain can be reduced by using a stereo viewer such as the Berezin ScreenScope, which features adjustable mirrors that reflect a stereo-image pair from any graphic display onto one's eyes, behind which the brain fuses the images into a single, sharp 3D scene. Since it uses mirrors, the ScreenScope supports full color, isn't haunted by ghost images, and is completely free from screen flicker, which allow for easy viewing of many sizes of images. The monitor's resolution determines the stereo image resolution.

Alternatively, binocular HMDs such as iO Display



(a) Schematic

(b) Juxtaposed Views

Figure 15. *Multiple synchronized monitors: turnoramic turnorama.*

Systems' i-glasses can also be used to present stereographic pairs to individuals. For projection methods suitable for display to groups of people, a dual display system can be configured, sending video signals for the left and right movies to separate projectors. These two projectors shine through two out-of-phase polarized filters onto the same place on a silver (not white) screen (such as that sold by Reel 3-D Enterprises) with retroreflective properties that preserve polarization, allowing users to enjoy stereo effects with compatibly polarized passive eyewear.

A multidisplay-configured computer can also drive a stereo projection system. Our university supports a 'reality center'-style 3DTheater which features such a system. A computer with three graphics ports (two on one card, one on another), configured as two logical displays by juxtaposing the dual-ported video signals into a single space, drives (through a splitter) both multiple monitors (for control and debugging) and also a two-channeled front-projection system. Our applications, including VR_4U_2C and Java3D-based Just Look at Yourself!, described in the next section, automatically fill these stereo frames, either explicitly (VR_4U_2C launches two separate windows that match each other's size) or implicitly (our Java3D applications have a single window with twin side-by-side panes that fill the lateral fields when the window is maximized, aligning the respective panes with the appropriate projection channel). About fifty guests can simultaneously enjoy such contents, and the CVE allows synchronous spatial audio through two collocated speaker array systems as well as control by mobile device (which has interfaces as illustrated in Figure 4).

4.3.3 Dolly- and Track-Enhanced SQTVR Browsing. The enhanced functionality described here has motivated extensions to the user interface. The conventional idiom for QTVR browsing is to interpret the ← and → arrow keys as panning imperatives (rotating the scene opposite subjective yaw— CW and CCW, respectively) and (for panos) the ↑ and ↓ arrows as tilt (climb and dive, respectively), while `Control` and `Shift` denote zoom out and in, respectively. As our track and dolly enhancement allows translation as well as rotation, the arrow keys have been extended: `Alt+` chorded combinations with keypad arrows invoke (transverse) sway and (longitudinal) surge, according to the natural planar interpretation.

5 Just Look At Yourself!: Visualization and Emulation via Java3D

Java3D, hereafter J3D, is a framework for dynamic virtual spaces. Our interface suite includes J3D stereoscopic perspective display and control clients, which can be configured at runtime to display multiple perspectives from various standpoints, including exocentrically from strategically placed cameras and egocentrically (endocentric and tethered) with respect to a selected humanoid avatar located between the nodes from which a panoramic stereo pair was captured. In particular, a left-right pair from any of these perspectives can be displayed in a multipaned window to display a scene stereographically.

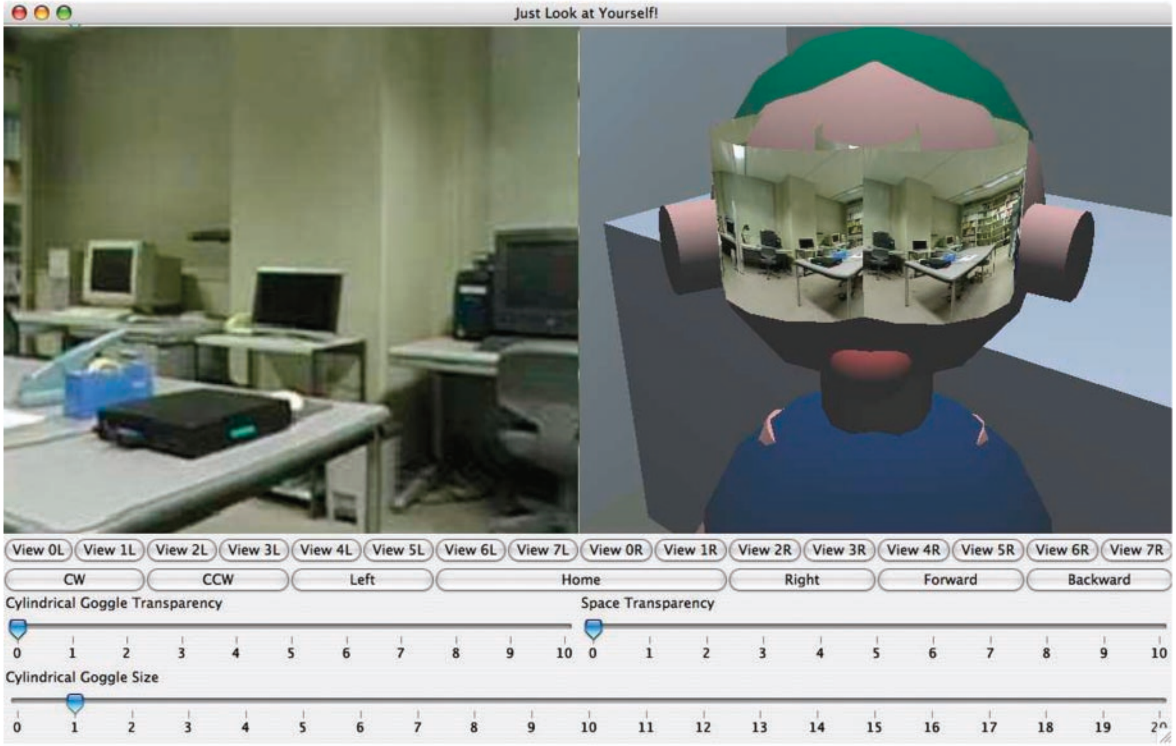
We use J3D to model the panoramic projection, including stereographic capability, through side-by-side image pairs. A humanoid in the scene, a figurative avatar, stands at the location corresponding to a pair of panoramic nodes. Cylinder pairs with texture maps corresponding to the viewpoint node are instantiated in the J3D scenegraph as goggles donned by the avatar, textured with the respective panoramic image, and centered at the eyes of an avatar at that node, as visible in Figure 16a right and 16b left. Back-face culling is disabled to use a single polygon with a double face (bifaceted), so the rendered texture map is also visible exocentrically. The respective cylinders, at the nodes shown in Figure 6, are activated and disactivated by dynamically setting/resetting the J3D `isVisible` node attribute according to which pair is active (depending on location and orientation of avatar).

The humanoid responds to locally generated and networked repositioning commands (simply quick sliding without realistic walking animation). The cylindrically texture-mapped panorama (dynamically selected) is coupled with the figurative avatar, but is rotation-invariant, since it is aligned with the space it portrays. J3D uses a tree-like hierarchical scene graph to model spaces, employing a dynamically parameterized `Transform3D` node to reposition descendent geometry. As diagrammed by Figure 17, by putting the cylinder and humanoid on sibling branches, they both inherit translational updates, but only the humanoid is subject to rotational commands.

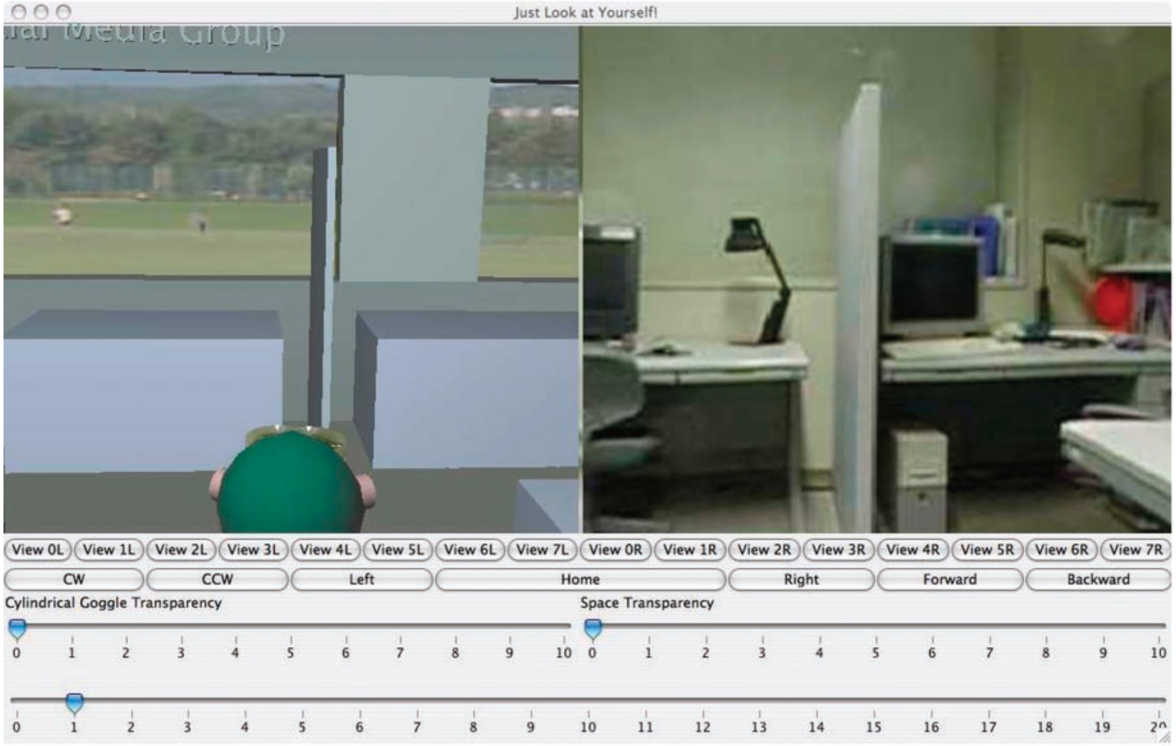
5.1 Virtual Cameras

Besides the endo- and egocentric perspectives, virtual cameras distributed around the space allow a variety of exocentric views, including plan (bird's-eye or God's-eye) and elevation (side). Putting J3D `ViewPlatforms` into both view and content branches of the scenegraph allows separate camera pairs, unattached to the moving objects, as well as inside or behind the head of the avatar. The perspective control (Hoeben & JanStapper, 2006) for Just Look at Yourself! is the same as that of VR_4U_2C, including all the 2.5D manipulations described in Section 4.3.3. 3D commands extend the controls: allowing the cameras to be barrel rolled ([top of head] left-right) and boomed (up-down).

To maximize flexibility, the J3D clients' window panes may be independently switched at runtime to display endo-, ego-, or exocentric perspectives, for respective eyes, so a complementary perspective selection can show a stereo pair, displayable via the techniques described earlier, switchable to mixed (security monitors) mode, a juxtaposition of independent non-stereographic perspectives. Phantom sources, invisible objects tagged with audio sources, extend J3D in a sibling client Multiplicity (also available from the aforementioned project web site) to allow separate listening and viewing positions by displaying egocentrically auditorily displayed sound sources reflecting exocentrically modeled and visually displayed soundscapes (Fernando, Adachi, & Cohen, 2004).



(a) Both panes are configurable to show an endo-, ego-, or exocentric stereographic IBR, sterographic CG rendering, or mixed perspective display.



(b) The sampled and synthesized scenes are aligned.

Figure 16. *Perspective displays by "Just Look at Yourself!" Java3D visualizer and emulator.*

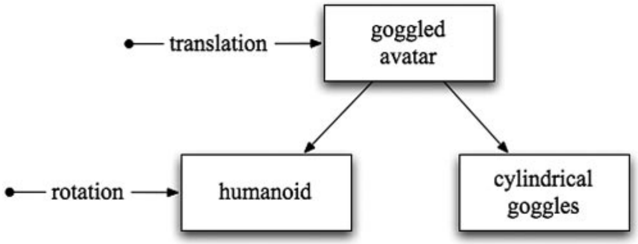


Figure 17. *Repositioning routing by scene graph branch.*

5.2 Dynamically Texture-Mapped Cylindrical Goggles

The transparency of the CG 3D space and also the photorealistic panoramic scene, as well as the size of the cylinder pair goggles upon which they are texture-mapped, can all be adjusted using sliders at runtime to understand the relationship of the projections. The cylinders are scaled and centered at the respective eyes vertically and horizontally, but we have not yet (because of some problems with J3D) been able to switch these textures dynamically. A full human field of view is about 120° vertically and 200° horizontally, the overlapped binocular field of view ranging from about 30° (Stanney, 2002, p. 31) to about 120° when focused at infinity (Salvendy, 1997, p. 1737). By setting VR_4U_2C's `Field of View` to 32° and specifying egocentric (`View 0`) perspective in Just Look at Yourself!, the interfaces' stereoscopic views are approximately identical. Except for quantization errors (as the track and dolly moves in discrete steps corresponding to the interocular distance) and level of detail (as a CAD model is coarser than a photograph of the real space), the (virtual) 3D spatial objects visible beyond the translucent goggles correspond to the image projected on the cylinder, as in Figure 16.

6 Conclusion

As panoramic and turnoramic browsers diffuse into practical applications, interfaces will be needed that combine egocentric and exocentric perspectives. The VR_4U_2C multimonitor and multidisplay QTVR browser, integrated with our heterogeneous groupware client suite, encourages multiperspective exploration. This approach is very affordable and can be used on desktop (or laptop) computers for fishtank VR as well as with theater projection for immersive experiences. Although stereo panoramic and turnoramic viewers have been developed by other groups, sometimes as multimonitor applications, to the best of our knowledge, ours is the first instance of track-enabled and dolly-enabled non-anaglyphic stereographic QuickTime VR (SQTVR). Of course such IBR techniques will leverage advances in computational photography (Bimber, 2006), such as dynamic or multifocal depth of field. This QTVR-based IBR egocentric interface and our J3D-based Just Look At Yourself! multiperspective interface complement each other. The VR_4U_2C client runs on Macs (via the Java Advanced Imaging package) and Windows PCs (there being no QuickTime for Java on Solaris or Linux as of this writing) but doesn't have any exocentric perspective; the J3D interface has a more flexible perspective and runs on those platforms as well as Sun workstations. Both clients support stereographic displays, so by selecting a first-person viewpoint, the J3D stereo rendering can be made to emulate the SQTVR panoramic browser. In that sense, exocentricism is a generalization of egocentricism.

Acknowledgments

We gratefully acknowledge the many useful suggestions of William L. Martens, as well as those of anonymous referees. The Java3D lab model was originally developed by a team led by Shiyūnō Kazuki and then modified by Tomoya Kamada, Hiroaki Osaka, and Takuzou Yoshikawa. Dr. Kouzo Watanabe provided the x-rays described in Section 1.4. This research was subsidized by a grant from the Fukushima Prefectural Foundation for the Advancement of Science.

References

Adachi, K., Cohen, M., Duminduwardena, U., & Kanno, K. (2004). "*Kuru-kuru* pitcher": A game for the ${}^{\mathrm{S}}_{\mathrm{c}}\mathrm{ha}_{\mathrm{i}}\mathrm{r}^{\mathrm{e}}$ internet chair. *Third International Conference on Entertainment Computing, ICEC,* Vol. 3166. *Lecture notes in computer science* (pp. 35–45). Berlin: Springer.

Adachi, K., Iwai, K., Yamada, E., & Cohen, M. (2005). Multimodal wayfinding in a driving simulator for the $^{S}_{c}ha_{i}r^{e}$ internet chair, a networked rotary motion platform. In F. Kishino, Y. Kitamura, H. Kato, & N. Nagata (Eds.). *ICEC,* Vol. 3711. *Lecture notes in computer science* (pp. 511–514). Berlin: Springer.

Adamson, C. (2005). *QuickTime for Java: A developer's notebook.* Sebastopol, CA: O'Reilly.

Aliaga, D. G., Funkhouser, T., Yanovsky, D., & Carlbom, I. (2003). Sea of images: A dense sampling approach for rendering large indoor environments. *IEEE Computer Graphics and Applications, 23*(6), 22–30.

Aliaga, D. G., Yanovsky, D., Funkhouser, T., & Carlbom, I. (2003). Interactive image-based rendering using feature globalization. *2003 Symposium on Interactive 3D Graphics, SI3D,* 163–170.

Amano, K., Matsushita, F., Yanagawa, H., Cohen, M., Herder, J., Martens, W., et al. (1998). A virtual reality sound system using room-related transfer functions delivered through a multispeaker array: The PSFC at the University of Aizu Multimedia Center. *Transactions of the Virtual Reality Society of Japan, 3*(1), 1–12.

Arijon, D. (1991). *Grammar of the film language.* Los Angeles: Silman-James Press.

ASCII Editing Group. (2001). *iMode Java programming.* Author.

Barrilleaux, J. (2001). *3D user interfaces with Java 3D.* Greenwich, CT: Manning Publications.

Bartoli, A., Dalal, N., & Horaud, R. (2004). Motion panoramas. *Computer Animation and Virtual Worlds, 15,* 501–517.

Benosman, R., & Kang, S. B. (2001). *Panoramic vision: Sensors, theory, and applications.* Berlin: Springer.

Bimber, O. (Ed.). (2006). Computational photography [Special issue]. *Computer, 39*(8).

Bolhassan, A., Cohen, M., & Martens, W. L. (2004). VR_4U_2C: A multiuser multiperspective panorama and turnorama browser using QuickTime VR and Java featuring multimonitor and stereographic display. *Transactions of the Virtual Reality Society of Japan, 9*(1), 69–78.

Bolhassan, N. A., Cohen, M., Fernando, O. N., Kamada, T., Martens, W. L., Osaka, H., et al. (2002). "Just Look at Yourself!": Stereographic exocentric visualization and emulation of stereographic panoramic dollying. *International Conference on Artificial Reality and Tele-Existence, ICAT,* 146–153.

Chan, Y.-F., Fok, M.-H., Fu, C.-W., Heng, P.-A., & Wong, T.-T. (1999). A panoramic-based walkthrough system using real photos. *The Seventh Pacific Conference on Computer Graphics and Applications,* PG' 99, 231.

Chen, S. E. (1995). QuickTime VR—An image-based approach to virtual environment navigation. *Computer Graphics, SIGGRAPH'95,* 29–38.

Chen, S. E., & Williams, L. (1993). View interpolation for image synthesis. *Computer Graphics, SIGGRAPH'93,* 279–288.

Chiang, C.-C., Way, D.-L., Shieh, J.-W., & Shen, L.-S. (1998). A new image morphing technique for smooth vista transitions in panoramic image-based virtual environment. *ACM Symposium on Virtual Reality Software and Technology, VRST'98,* 81–90.

Cohen, M. (1994). Cybertokyo: A survey of public VRtractions. *Presence: Teleoperators and Virtual Environments, 3*(1), 87–93.

Cohen, M. (1998). Quantity of presence: Beyond person, number, and pronouns. In T. L. Kunii & A. Luciani (Eds.). *Cyberworlds* (pp. 289–308). Berlin: Springer.

Cohen, M. (2002). Rotational degrees-of-freedom QTVR object movie. Retrieved April 2, 2002 from sonic.u-aizu.ac.jp/spatial-media/QTVR/Rotational-DsoF.mov

Cohen, M. (2003). The internet chair. *International Journal of Human Computer Interaction (Special Issue: Mediated Reality), 15*(2), 297–311.

Cohen, M. (2005a). Shoe QTVR object movie (turnorama). Retrieved December 2, 2005 from sonic.u-aizu.ac.jp/spatial-media/QTVR/shoe.mov

Cohen, M. (2005b). Dental QTVR object movie-style pano. Retrieved December 2, 2005 from sonic.u-aizu.ac.jp/spatial-media/QTVR/teeth.mov

Cohen, M., Azumi, T., Yatsuyanagi, Y., Sasaki, M., Yamaoka, S., & Takeichi, O. (2002). Networked speaker array streaming back to client: The world's most expensive sound spatializer? *International Conference on Artificial Reality and Tele-existence, ICAT,* 162–169.

Cohen, M., & Sasa, K. (2000). An interface for a soundscape-stabilized spiral-spring swivel-seat. *The Seventh Western Pacific Regional Acoustics Conference, WESTPRAC VII,* 321–324.

Cortázar, J., & Antonioni, M. (1966). Blowup [Motion picture]. United States: Metro-Goldwyn-Mayer.

Cruz-Neira, C., Sandin, D., & DeFanti, T. (1993). Surround-screen projection-based virtual reality: Design and implementation of the CAVE. *Computer Graphics, SIGGRAPH'93,* 135–142.

Darsa, L., Silva, B. C., & Varshney, A. (1997). Navigating static environments using image-space simplification and morphing. *Symposium on Interactive 3D Graphics,* 25–34.

Davis, E. T., & Hodges, L. F. (1995). Human stereopsis, fusion, and stereoscopic virtual environments. In W. Barfield & T. A. Furness, III (Eds.), *Virtual Environments and Advanced Interface Design* (pp. 145–174). Oxford, UK: Oxford University Press.

Dick, P. K., & Scott, R. (1982). Blade runner [Motion picture]. United States: Warner Brothers.

Diner, D. B., & Fender, D. H. (1993). *Human engineering in stereoscopic viewing devices.* New York: Plenum Press.

Drouin, F. (1995). *The stereoscope and stereoscopic photography.* Culver City, CA: Reel 3-D Enterprises. (Original work published 1894).

Duminduwardena, U. C., & Cohen, M. (2004). Controlling the $^{S}_{c}\text{ha}_{i}\text{r}^{e}$ internet chair with a mobile device. *Fourth International Conference on Computer Information Technology,* 215–220.

Favalora, G. E. (2005). Volumetric 3D displays and application infrastructure. *Computer, 38*(8), 37–45.

Feng, Y., & Zhu, J. (2001). *Wireless Java programming with J2ME.* Indianapolis, IN: Sams.

Fernando, O. N. N. (2006). A multiuser multiperspective stereographic QTVR browser complemented by Java3D visualizer and emulator. Retrieved June 15, 2006 from sonic.u-aizu.ac.jp/spatial-media/Mixed-Reality.html

Fernando, O. N. N., Adachi, K., & Cohen, M. (2004). Phantom sources for separation of listening and viewing positions for multipresent avatars in narrowcasting collaborative virtual environments. *International Workshop on Multimedia Systems and Applications,* 170–175.

Fernando, O. N. N., Adachi, K., Duminduwardena, U., Kawaguchi, M., & Cohen, M. (2006). Audio narrowcasting and privacy for multipresent avatars on workstations and mobile phones. *IEICE Transactions on Information and Systems, E89-D*(1), 73–87.

Fernando, O. N. N., Cohen, M., Duminduwardena, U. C., & Kawaguchi, M. (2006). Duplex narrowcasting operations for multipresent groupware avatars on mobile devices. *International Journal of Wireless and Mobile Computing (Special Issue on Mobile Multimedia Systems and Applications), 1*(5).

Fernando, O. N. N., Saito, G., Duminduwardena, U., Tanno, Y., & Cohen, M. (2005). Cloning and teleporting avatars across workstations and mobile devices in collaborative virtual environments: Clipboard operations for virtual reality. *International Conference on Information and Automation.*

Fu, C.-W., & Heng, P.-A. (1998). A fast algorithm to achieve perceptually acceptable views during multi-node panoramic walk-through. *Workshop on Computer Graphics,* 35–39.

Fu, C.-W., Wong, T.-T., & Heng, P.-A. (1998). Triangle-based view interpolation without depth-buffering. *ACM Journal of Graphics Tools, 3*(4), 13–31.

Fu, C.-W., Wong, T.-T., & Heng, P.-A. (1999). Computing visibility for triangulated panoramas. *Tenth Eurographics Workshop on Rendering,* 169–182.

Fukuoka, T. (Ed.). (2004). *Beginning J2ME: From novice to professional.* Tokyo: ASCII.

Gordon, R., & Talley, S. (1999). *Essential JMF—Java media framework.* Englewood Cliffs, NJ: Prentice Hall.

Hayashi, R., Nakaizumi, F., Yano, H., & Iwata, H. (2005). Development of a full-surround stereo spherical immersive display using multiple projectors. *Transactions of the Virtual Reality Society of Japan, 10*(2), 163–172.

Hirose, M. (1997). Image-based virtual world generation. *IEEE Multimedia, 4*(1), 27–33.

Hirose, M., Watanabe, S., & Endo, T. (1998). Generation of wide-range virtual spaces using photographic images. *The Fourth IEEE Virtual Reality Annual International Symposium, VRAIS'98,* 234–241.

Hirouchi, T. (2004). *Java 3D graphics.* Tokyo: Serendip. (In Japanese).

Hoeben, A., & JanStapper, P. (2006). Taking clues from the world outside: Navigating interactive panoramas. *Personal and Ubiquitous Computing, 10*(2–3), 122–127.

Horstmann, C. S., & Cornell, G. (2001). *Core Java 2, Volume I—Fundamentals.* Englewood Cliffs, NJ: Prentice Hall.

Huang, H.-C., & Hung, Y.-P. (1998). Panoramic stereo imaging system with automatic disparity warping and seaming. *Graphical Models Image Processing, 60*(3), 196–208.

Ishikawa, S., Saito, Y., & Cohen, M. (2005). Synchronization of networked avatars and appliances with mobile phone ringtones: Integrating Java3D and LAN-tap roomware with J2ME. *Fifth International Conference on Computer and Information Technology, CIT,* 553–557.

Jacobs, C. (2004). *Interactive panoramas: Techniques for digital panoramic photography.* Berlin: Springer.

Jorke, H., & Fritz, M. (2003). A new stereoscopic visualization tool by wavelength multiplex imaging. *Proceedings of Electronic Displays.*

Kanno, K., Fernando, N., Bolhassan, A., Narita, S., & Cohen, M. (2006). "Personal practically panoramic" multimodal interfaces. *IEEE VR'06 Workshop on Emerging Display Technologies, EDT,* 37–40.

Kanno, T., Cohen, M., Nagashima, Y., & Hoshino, T. (2001). Mobile control of multimodal groupware in a distributed virtual environment. *International Conference on Artificial Reality and Tele-existence, ICAT,* 147–154.

Katayama, A., Tanaka, K., Oshino, T., & Tamura, H. (1995). A viewpoint dependent stereoscopic display using interpolation of multi-viewpoint images. *Stereoscopic Displays and Virtual Reality Systems II, SPIE, 2409,* 11–20.

Kawakita, Y., Hamaguchi, Y., Tsukamoto, A., & Miyazaki, T. (2000). Generation of panoramic stereo images from movie using single video camera. *Journal of the Information Processing Society of Japan, 41*(6), 1848–1856.

Kitchens, S. A. (1998). *The QuickTime VR book: Creating immersive imaging on your desktop.* Berkeley, CA: Peachpit Press.

Knudsen, J. (2003). *Wireless Java: Developing with J2ME* (2nd ed.). Berkeley, CA: Apress.

Koizumi, N., Cohen, M., & Aoki, S. (2000). Japanese patent #3042731: Sound reproduction system.

Kontio, M. (2003). *Mobile Java with J2ME.* Edita, Finland: IT Press.

Kotake, D., Endo, T., Pighin, F., Katayama, A., Tamura, H., & Hirose, H. (2001). Cybercity walker 2001: Walking through and looking around a realistic cyberspace reconstructed from the physical world. *International Symposium on Mixed Reality,* 205–206.

Lea, D. (2000). *Concurrent Programming in Java* (2nd ed.). Addison-Wesley.

Li, S., & Knudsen, J. (2005). *Beginning J2ME: From novice to professional* (3rd ed.). Berkeley, CA: Apress.

Lippman, A. (1980). Movie-maps: An application of the optical videodisc to computer graphics. *Computer Graphics, SIGGRAPH'80,* 32–42.

Maeda, H., Tanikawa, T., Yamashita, J., Hirota, K., & Hirose, M. (2004). Real world video avatar: Transmission and presentation of human figure. *IEEE Virtual Reality Conference,* 237–238.

Mahmoud, Q. H. (2002). *Wireless Java.* Sebastopol, CA: O'Reilly.

Maremaa, T., & Stewart, W. (2005). *QuickTime for Java: A programmer's guide to building multimedia applications with Java* (2nd ed.). San Mateo, CA: Morgan Kaufmann.

Martens, W. L., McRuer, B., Childs, C. T., & Virree, E. (1996). Physiological approach to optimal stereographic game programming: A technical guide. *Stereoscopic Displays and Virtual Reality Systems III, IS&T/SPIE, 2653,* 261–270.

Matusik, W., & Pfister, H. (2004). 3D TV: A scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Transactions on Graphics, 23*(3), 814–824.

McAllister, D. F. (Ed.). (1993). *Stereo computer graphics and other true 3D technologies.* Princeton, NJ: Princeton University Press.

McMillan, L., & Bishop, G. (1995a). Plenoptic modeling: An image-based rendering system. *Twenty-second Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH'95,* 39–46.

McMillan, L., & Bishop, G. (1995b). Plenoptic modeling: An image-based rendering system. *Computer Graphics, 29,* 39–46.

Mikuriya, T., Shimizu, M., & Cohen, M. (2001). A collaborative virtual environment featuring multimodal information controlled by a dynamic map. *3D Forum: Journal of Three Dimensional Images, 15*(1), 133–136.

Milgram, P., & Coquhoun, H., Jr. (1999). A taxonomy of real and virtual world display integration. In Y. Ohta & H. Tamura (Eds.), *Mixed reality: Merging real and virtual worlds* (pp. 5–30). Berlin: Springer.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems (Special Issue on Networked Reality), E77-D*(12), 1321–1329.

Mohr, M. (2004). Spatial narrative "The Island of Misfit Toys." *Computer Graphics, SIGGRAPH'04.*

Naemura, T., Kaneko, M., & Harashima, H. (1998). Multiuser immersive stereo in projection-based virtual reality system. *Third Virtual Reality Society of Japan,* 57–60.

Nagai, T., Cohen, M., Moriguchi, Y., & Murakami, Y., (2007). A dual-driver system of networked fire truck simulator with rotating motion platform and force display. IEEE TICE 2007: *Tangible Interaction in Collaborative Environments.*

Nagata, S. (2002). Three-dimensional and autostereoscopic image displays toward unconstrained and interactive viewing. *Journal of the Institute of Electronics, Information and Communication Engineers, 85*(1), 43–48. (In Japanese).

Naimark, M. (1997). A 3D moviemap and a 3D panorama. *SPIE/SPSE Electronic Imaging, 2012.*

Nalwa, V. S. (2000a). Panoramic viewing system with offset virtual optical centers. U.S. Patent 6,111,702.

Nalwa, V. S. (2000b). Stereo panoramic viewing system. U.S. Patent 6,141,145.

Nayar, S. K. (1997). Catadioptric omnidirectional camera. *IEEE Computer Vision and Pattern Recognition (CVPR).*

Oliveira, M. M., & Bishop, G. (1999). Image-based objects. *ACM Symposium on Interactive 3D Graphics,* 191–198.

Ota, A. (2003). *Yasasiku Wakaru Java 3D*. Software Technology. Tokyo: Gijutsu. (In Japanese).

Palmer, I. (2001). *Essential Java 3D fast: Developing 3D graphics applications*. Berlin: Springer.

Pausch, R., Burnette, T., Brockway, D., & Weiblen, M. E. (1995). Locomotion in virtual worlds via flight into hand-held miniatures. *Computer Graphics, SIGGRAPH'95,* 399–400.

Peleg, S., & Ben-Ezra, M. (1999). Stereo panorama with a single camera. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* 395–401.

Peleg, S., Ben-Ezra, M., & Pritch, Y. (2001). Omnistereo: Panoramic stereo imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23*(3), 279–290.

Pritch, Y., Ben-Ezra, M., & Peleg, S. (2000). Automatic disparity control in stereo panoramas (OmniStereo). *IEEE Workshop on Omnidirectional Vision, OMNIVIS00,* 54–61.

Rademacher, P., & Bishop, G. (1998). Multiple-center-of-projection images. *Computer Graphics, SIGGRAPH'98,* 199–206.

Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S., & Forlines, C. (2003). iLamps: Geometrically aware and self-configuring projectors. *ACM Transaction Graphics, 22*(3), 809–818.

Raskar, R., van Baar, J., Willwacher, T., & Rao, S. (2004). Quadric transfer for immersive curved screen displays. *Computer Graphics Forum, 23*(3), 451–460.

Sakagawa, Y., Katayama, A., Kotake, D., & Tamura, H. (2002). Data compression and hardware implementation of ray-space rendering for interactive augmented virtuality. *Presence: Teleoperators and Virtual Environments, 11*(2), 203–219.

Salvendy, G. (Ed.). (1997). *Handbook of human factors and ergonomics,* 2nd ed. New York: Wiley.

Sasaki, M., & Cohen, M. (2004). Dancing music: Integrated MIDI-driven synthesis and spatialization for virtual reality. *Proceedings of the Audio Engineering Society Conference, AES,* Preprint 6316 (R-3).

Seitz, S. M., & Dyer, C. R. (1996). View morphing. *Computer Graphics, SIGGRAPH'96,* 21–30.

Selman, D. (2002). *Java 3D programming*. Greenwich, CT: Manning.

Shade, J. W., Gortler, S. J., He, L.-W., & Szeliski, R. (1998). Layered depth images. *Computer Graphics, SIGGRAPH'98,* 231–242.

Shimada, D., Tanahashi, H., Kato, K., & Yamamoto, K. (2001). Extract and display moving object in all direction by using stereo omnidirectional system (SOS). *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, 3DIM,* 42–47.

Shum, H.-Y., Kalai, A., & Seitz, S. M. (1999). Omnivergent stereo. *The Seventh International Conference on Computer Vision,* 22–29.

Snavely, N., Seitz, S. M., & Szeliski, R. (2006). Photo tourism: Exploring photo collections in 3D. *Transactions on Graphics, SIGGRAPH'06, 25,* 835–846.

Sowizral, H., Rushforth, K., & Deering, M. (1998). *The Java 3D API specification (Japanese edition)*. Reading, MA: Addison-Wesley.

Sowizral, H., Rushforth, K., & Deering, M. (2000). *The Java 3D API specification* (2nd ed.). Reading, MA: Addison-Wesley.

Stanney, K. M. (Ed.). (2002). *Handbook of virtual environments: Design, implementation, and applications. Human Factors and Ergonomics*. Mahwah, NJ: Lawrence Erlbaum.

Stoakley, R., Conway, M. J., & Pausch, R. (1995). Virtual reality on a WIM: Interactive worlds in miniature. *SIGCHI Conference on Human Factors in Computing Systems,* 265–272.

Szeliski, R., & Shum, H.-Y. (1997). Creating full view panoramic image mosaics and environment maps. *Computer Graphics, SIGGRAPH'97,* 251–258.

Tanaka, K., Hayashi, J., Endo, T., & Tachi, S. (2003). A method for panoramic stereo image acquisition. *ICAT,* 257–262.

Tanaka, S., Matsubara, S., Monobe, K., Miura, S., Miyawaki, Y., Hattori, H., et al. (2002). *Java3D graphics introduction*. Tokyo: Morikita. (In Japanese).

Topley, K. (2002). *J2ME in a Nutshell.* Sebastopol, CA: O'Reilly.

Tzavidas, S., & Katsaggelos, A. K. (2005). A multicamera setup for generating stereo panoramic video. *IEEE Transactions on Multimedia, 7*(5), 880–890.

Vacca, J. R. (2002). *I-Mode crash course.* New York: McGraw-Hill.

Vineyard, J. (2000). *Setting up your shots: Great camera moves every filmmaker should know*. Studio City, CA: Michael Wiese Productions.

Waldsmith, J. (2002). *Digital image processing* (2nd ed.). Iola, WI: Krause Publications.

Walrath, K., & Campione, M. (1999). *The JFC swing tutorial.* Reading, MA: Addison-Wesley.

Walsh, A. E., & Gehringer, D. (2002). *Java 3D API jump-start*. Englewood Cliffs, NJ: Prentice Hall.

Watkinson, J. (2001). *Convergence in broadcast and communications media*. Oxford, UK: Focal Press.

Yamazaki, Y. (2001). *How to make an iAppli.* Tokyo: SoftBank.

Yendo, T., Kawakami, N., & Tachi, S. (2005). Seelinder: The cylindrical lightfield display. *ACM SIGGRAPH 2005 Emerging Technologies,* 16.

Zheng, J. Y. (2003). Digital route panoramas. *IEEE Multimedia, 10*(3), 57–67.