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

Michael Cohen, Noor Alamshah Bolhassan, and Owen Noel Newton Fernando

Spatial Media Group, Computer Arts Lab.

University of Aizu

Aizu-Wakamatsu, Fukushima-ken 965-8580

Japan

mcohen@u-aizu.ac.jp

voice: [+81](242)37-2537

fax: [+81](242)37-2772

Second author's current address:

Faculty of Computer Science and Information Technology

Universiti Malaysia Sarawak

94300 Kota Samarahan

Sarawak, Malaysia

bnalamshah@fit.unimas.my

Third author's current address:

Mixed Reality Lab., Dept. of Electrical and Computer Engineering

National University of Singapore

7 Engineering Drive 1, Blk E3A, #02-04/05

Singapore 117574

newtonfernando@mixedrealitylab.org

Note the use of emboldening and small caps, as in 2D, 2.5D, 3D, 3D-CGI, Java3D) to denote acronyms: (CAD, CAT [**c**omputer **a**xial **t**omography], CMOS, GIF, MIDlets) and initials AR [**a**ugmented **r**eality], (CDR [**c**omputed **d**ental **r**adiography], CG, CGI, CPU, **c**ollaborative **v**irtual **e**nvironments (CVEs), CW and CCW, DV [**d**igital **v**ideo], "**f**irst-**p**erson **s**hooter" (FPS), HMDs, IBR [**i**mage-**b**ased **r**endering], J3D, JME, JMF, JSE, **l**iquid-**c**rystal **d**isplay (LCD), MR [**m**ixed **r**eality], QTVR, SQTVR. Also note the exotic orthography of VR_4U_2C, the designation of keys as in `Control`, `Shift`, & Alt+, ←, →, ↑, & ↓; the use of monospacing to indicate computer modes and keywords, `anaglyphic`, `parallel`; `CVEClientIF` , `get` and `set`; `parallel`, `cross-eyed`, `over/under`; `point 1` etc.; "`View 0`" etc. Also note copious use of footnotes of URLs for commercial web pages, for which including the access date (as suggested by the Presence style guide) would be inappropriate.

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, "$VR_4 U_2 C$" ('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* **s***tereographic* 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 Java3*D *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 $VR_4 U_2 C$, 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 former is conceptually viewed and upon which the latter are texture-mapped, can be adjusted at runtime to understand the relationship of the spaces.

1. Introduction

1.1. Points of View

A classic example of an exocentric display is a map. But if one allows themself 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 et al., 1995; Stoakley et al., 1995).

Figure 1 here.

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

Float (a.k.a. "Predict") (Fig. 1(c)), 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") (Fig. 1(d)), in which the monitor tracks the car as if from a blimp, clearly showing one's own car in the context of the field.

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 (**c**omputer **g**raphic) 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) (Cohen, 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 1st- and 2nd-person metaphors. We reserve the neologism "endocentric" ('centered within') for intimate, strictly 1st-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 **c**omputer-**g**enerated **i**magery (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 straight-forward methodology, like mesh-generating CAD (**c**omputer-**a**ided **d**esign), is often inappropriate for

Perspective		Person	Proximity	Objectification	Virtua Racing	QTVR
egocentric	endocentric	1st-	proximal	reflexive	Cockpit (Drive or Dynamic)	panorama
egocentric	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

Table 1. Continuum of Navigation Modes

	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

Table 2. Polygon vs. Image-Based Rendering

geometric modeling of large-scale and complex areas like a cityscape. An IBR (**i**mage-**b**ased **r**endering) approach— including techniques such as HoloMedia, Light Field, Lumigraph, Delta Tree, and ray-space methods (Sakagawa et al., 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 trade-offs 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 "$\mathrm{VR_4U_2C}$" 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 pre-recorded samples, as by the various synthetic singers based upon the Yamaha Vocaloid.[1] 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.

Figure 2 here.

[1] www.vocaloid.com, www.zero-g.co.uk/index.cfm?articleid=804

1.3 Augmented and Mixed Reality

The basic idea of **a**ugmented **r**eality (AR) or **m**ixed **r**eality (MR) in graphics is spatial sensation via 2D information upon 3D models (Milgram & Kishino, 1994; Milgram & Coquhoun, Jr., 1999). The "virtuality continuum" observes that the real/virtual dichotomy is not sharp, but interpolatively smooth. Idealised 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[2]) (Chen, 1995) is one of many computer-based interactive image-based technologies, presenting both panoramas[3] (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

[2] `apple.com/quicktime/technologies/qtvr`
[3] `panoguide.com`

		Representation	
		Cyclorama or Panorama (single image)	Turnorama (array of frames)
Experience	Outward-looking	ordinary **pano**rama	panorama as turnorama
	Inward-looking	turnorama as panorama: unnatural	ordinary **turno**rama
Rotation		subject	object

Table 3. QTVR Representation × Experience

rotating object from a static viewpoint— is represented as an array of images. A cyclorama[4] 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 15(b)). Table 4 reviews the pertinent spatial dimensions and juxtaposes cinematographic correspondences (Arijon, 1991) (Vineyard, 2000).

Post-production authoring tools[5] (including VR Worx[6]) 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 in Boston[7], 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[8] with digital video camera interface can be used to capture a sequence of stills to make a turnorama.[9] Normally such capturing and experiences are exocentric, but we put the camera *on the turntable* to

[4]www.acmi.net.au/AIC/CYCLORAMA.html
[5]worldserver.com/turk/quicktimevr/QTVRlinks.html
[6]vrtoolbox.com
[7]http://www.marybakereddylibrary.org/exhibits/mapparium.jhtml
[8]spinimagedv.com
[9]sonic.u-aizu.ac.jp/spatial-media/QTVR/shoe.mov, retrieved December 2, 2005

Position

Static (Posture) / **Dynamic (Gesture)**

Location (Displacement)	**Scalar**	**Translation: Camera Motion**	**Directions (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

Table 4. Physically Spatial Dimensions: taxonomy of positional degrees of freedom, including cinematographic gestures. (Axis descriptions assume right-handed coordinate system with z gravitational up.)

capture an animated panoramic sequence.[10] (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[11] is a dental x-ray shot (literally) using a CDR (**c**omputed **d**ental **r**adiography) 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 (**c**omputer **a**xial **t**omography), 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 et al., 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

[10] `sonic.u-aizu.ac.jp/spatial-media/QTVR/Rotational-DsoF.mov`, retrieved April 2, 2002

[11] `sonic.u-aizu.ac.jp/spatial-media/QTVR/teeth.mov`, retrieved December 2, 2005

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 inches], 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 et al., 1995) (Shade et al., 1998) (Rademacher & Bishop, 1998) (Oliveira & Bishop, 1999) (Snavely et al., 2006).[12]

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 (**g**raphics **p**rocessing **u**nits, 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

[12] `labs.live.com/photosynth/`, `phototour.cs.washington.edu/`

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,[13] 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 subsection of this introduction describes the context and motivation for the development of this **s**tereographic 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

[13] developer.apple.com/documentation/QuickTime/RM/Fundamentals/QTOverview/QTOverview.pdf

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, “$\mathrm{VR_4U_2C}$” (‘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 counter-example, 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 straight-forward approach recalls the classic “Movie Map” (Lippman, 1980) (Naimark, 1997). Like the “Sea of Images” approach (Aliaga et al., 2003a; Aliaga et al., 2003b), 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 & 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 "**f**irst-**p**erson **s**hooter" (FPS) games recreate an actual place (like 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 trade-off is that a photogrammetric process requires a skilled modeller and provides continuous viewing, whereas our SQTVR capture process is simpler but more tedious, and yields pure stereographic pairs.

Rather than mosaicking captured video (Bartoli et al., 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 & 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 et al., 2004), or edge-blending techniques, like that used by (Hayashi et al., 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[14] and panoramic capture,[15] stitching,[16] and display,[17] including Cave-like (Cruz-Neira et al., 1993) **s**patially **i**mmersive **d**isplays[18] and artistic installations like (Mohr, 2004). Advanced and expensive equipment, such as Panoscan[19] or FullView[20] (Nalwa, 2000a; Nalwa, 2000b) cameras, can be used to get very high-resolution panoramic images, and more affordable solutions have also been commercialized, like the Eizoh Hyperomni Mirror system.[21] Users may also prefer to use either a slit-scan panorama camera[22] or coupled multiple cameras (Tzavidas & Katsaggelos, 2005).[23] Currently commercially available display systems include the Elumens[24] VisionDome and VisionStation, the Matsushita Panasonic Shiodome CyberDome[25] and CyberDome 1800, and the SpinDome.[26] Toshiba is developing a personal (6 pound) head-tracked viewing dome helmet with a 16 inch 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

[14] loreo.com, stereovisioninc.com/3d_vucam.html, tdvision.com

[15] 0-360.com, behere.com, channel360.com, fullview.com, imoveinc.com/geoview.php, ptgrey.com/products/ladybug2

[16] easypano.com, panoramafactory.com, panoramas.dk/panorama, ptgui.com, realviz.com, remotereality.com, www.cis.upenn.edu/~kostas/omni.html

[17] asuna-3d.com, eonreality.com, go-l.com/monitors/grand_canyon/features, imatronics.com, kinoma.com/products.html?player3, openVR.de, panoramtech.com, seamlessdisplay.com, seereal.com, vreye.jp, vri.ca

[18] www.fakespace.com

[19] panoscan.com

[20] fullview.com

[21] eizoh.net

[22] spheron.com

[23] www.imoveinc.com, genextech.com, immersivemedia.com

[24] elumens.com

[25] virtools.com/applications/learning-matsushita.asp

[26] solidray.co.jp/product/eizou/spin/spin_b.html

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 & design, simulation & training, entertainment, education, and medical research.

One of many advanced systems for volumetric display (Favalora, 2005) is the Perspecta Spatial 3D,[27] a 20-inch dome displaying full-color and -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 et al., 2005), a cylindrical three-dimensional color television, and the Spinning Display (Maeda et al., 2004), a flat **l**iquid-**c**rystal **d**isplay (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, etc.— the multi-window and -monitor 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.[28] A traditional technique (Huang & Hung, 1998) uses a rotating stereo 'head' with two displaced cameras,

[27] actuality-systems.com
[28] astronomy.swin.edu.au/~pbourke/stereographics/stereopanoramic/, www.lkn.ei.tum.de/~ingob/IBR/SP.htm, www.stereomaker.net/panorama/panoramae.htm

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 (S.O.S.) (Shimada et al., 2001),[29] which has sixty color CMOS cameras and omnidirectionally captures raw and depth images synchronized in realtime. Of course such an array is prohibitively expensive for general users.

Alternatively, a more convenient technique (Naemura et al., 1998) (Peleg & Ben-Ezra, 1999) (Kawakita et al., 2000) uses video cameras to prepare a stereo panorama by mosaicking 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,"[30] proposes a special camera catadioptric (reflection and refraction via mirror and lens) assembly, including a spiral lens or mirror (Shum et al., 1999) (Pritch et al., 2000) (Peleg et al., 2001) or a catacaustic of a cylindrical (or cylindroidal) mirror (Tanaka et al., 2003). The BeNoGo[31] 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[32] (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

[29] viewplus.co.jp/products/sos/astro-e.html
[30] humaneyes.com
[31] benogo.dk
[32] vstone.co.jp/top/products/sensor/

(Szeliski & Shum, 1997).

2.2.2. Navigation/Walkthroughs

Current image-based panoramic visualization browsers[33] 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 et al., 1998), mesh triangulation and morphing (Darsa et al., 1997), epipolar geometry (Chiang et al., 1998), and pixel- (McMillan & Bishop, 1995b) and triangle-based (Fu et al., 1998; Fu & Heng, 1998; Fu et al., 1999) (Chan et al., 1999) image warping techniques, which cause a source image to melt, dissolve, and re-arrange 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

[33] `apple.com/quicktime/qtvr`, `ipix.com`, `vrex.com`, `webuser.hs-furtwangen.de/~dersch/PTVJ/doc.html`, `zoomify.com`

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[34] and turnoramas[35] 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[36]) 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[37] 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[38] (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.

---

[34] 209.196.177.41/12/12-07.htm, museum.state.il.us/mic_home/3d/index.html
[35] www.texnai.co.jp/eng/stereo/stereo_top.html
[36] colorcode3d.com
[37] mpfwww.jpl.nasa.gov/MPF/vrml/qtvr_stereo.html
[38] infitec.net

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 et al., 2001) (Fernando et al., 2006) to support **c**ollaborative **v**irtual **e**nvironments (CVEs), allowing distributed users to share multimodal virtual spaces. Our CVE architecture is based upon a simple **c**lient/**s**erver 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,[39] including JSE [**J**ava **S**tandard **E**dition: core/desktop] (Lea, 2000) (Horstmann & Cornell, 2001), Java Sound, JMF [**J**ava **M**edia **F**ramework] (Gordon & Talley, 1999) and its media API rival 'QuickTime for Java' (Maremaa & Stewart, 2005) (Adamson, 2005), Java3D (Sowizral et al., 1998; Sowizral et al., 2000) (Palmer, 2001) (Tanaka et al., 2002) (Walsh & Gehringer, 2002) (Selman, 2002) (Ota, 2003) (Hirouchi, 2004), JME [**J**ava **M**icro **E**dition: 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

$\mathbf{VR_4U_2C}$ (Bolhassan et al., 2004),

[39] `java.sun.com`

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 $^{\mathrm{S}}_{\mathrm{c}}\mathbf{ha_{i}r^{e}}$ (for 'shared chair') **Internet Chair** (Koizumi et al., 2000) (Cohen, 2003) (Duminduwardena & Cohen, 2004) (Kanno et al., 2006) rotary motion platform (shown in Figure 3),

2.5D **Dynamic Maps** (Mikuriya et al., 2001),

the Pioneer Sound Field Controller (PSFC) proxy (Amano et al., 1998) enabling a 15-channel speaker array,

Java3D **widgets** for games (Adachi et al., 2004) and simulators (Adachi et al., 2005) (Nagai et al., 2006), and

mobile applications, including these MIDlets (JME **m**obile **i**nformation **d**evice applets) for mobile phones

"μVR$_4$U$_2$C" shown in Figure 4(a), &

"$\iota\cdot$Con" (Kanno et al., 2001) (Ishikawa et al., 2005) (Fernando et al., 2005) (Fernando et al., 2007) narrowcasting interface, shown in Figure 4(b),

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

Figure 3 here.

Figure 4 here.

The juxtaposition of these stereographic browsers, is this focus of this article. "VR_4U_2C" is a photorealistic, endocentric (1^{st}-person), image-based browser allowing panoramic panning, as well as tilting and zooming. "Just Look at Yourself!" is a **c**omputer **g**raphic browser, allowing more flexible camera positions, including endocentric (1^{st}-person: from the point-of-view of the avatar), tethered (2^{nd}-person: attached to but separate from the avatar), and exocentric (3^{rd}-person: totally detached from the avatar) perspectives (Barrilleaux, 2001, p. 125–147). They are respectively described in the following sections. Interested readers are encouraged to visit the project web page[40] 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. "VR_4U_2C": Networked Multimonitor SQTVR with Dollying and Tracking

4.1. Stereographic Panorama and Turnorama Imaging

4.1.1. Monocular Panoramic Capture

Figure 5 here.

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[41] and EyeSee360[42] 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

[40] sonic.u-aizu.ac.jp/spatial-media/Mixed-Reality.html
[41] kaidan.com
[42] EyeSee360.com

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[43] 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. Firstly, 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.

Figure 6 here.

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

[43] `EyeSee360.com/photowarp`

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.

Figure 7 here.

The numbered nodes in Figure 6(b) 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°), $\mathrm{VR_4U_2C}$ 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, $\mathrm{VR_4U_2C}$ monitors the azimuth and changes the right-eyed panorama to the appropriate node as the angle reaches a certain threshold. If the user dollys to the right, the left-eye panorama will change to `node 2` and the right-eye panorama to `node 8`; if the user dollys 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.

Figure 8 here.

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 $^{\mathrm{S}}_{\mathrm{c}}\mathrm{ha_ir^e}$ Internet Chair has significant sluggishness, a consequence of mechanical inertia and user comfort. When its software controller receives a target azimuth, it sends continuous updates on a separate channel whilst twisting the motion platform towards 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 aysnchronous updates of the lateral panes, and the stereo stitching is not entirely seamless; there are subtle but distinct “hick-coughs” 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, p. 382–3). 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 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

Figure 9 here.

As mentioned earlier in the Introduction's § 1.4, turnoramas (object movies) can be photographed using, for example, a FireWire-enabled digital video camera, a modern computer, and 3-D Object Imaging Kit[44] (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.

Figure 10 here.

Users can browse QTVR turnos stereoscopically (as shown in Figure 10) through $\mathrm{VR_4U_2C}$'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.

Figure 11 here.

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. $\mathrm{VR_4U_2C}$ 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

[44] `kaidan.com/products/DVkit.html`

shrinking of the perceived world, as the user unconsciously recalibrates perception to the natural interocular distance. (However, due to 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, $\mathrm{VR_4U_2C}$ connects to a session server, as previously described in § 3, to synchronously exchange parameters with sibling clients, including any other instances of $\mathrm{VR_4U_2C}$. 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, $\mathrm{VR_4U_2C}$ 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.) $\mathrm{VR_4U_2C}$ also gets zoom values and left & 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

Figure 12 here.

With the multiple synchronized displays feature, users may use $\mathrm{VR_4U_2C}$ 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,[45] 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.

Figure 13 here.

Figure 14 here.

Figure 15 here.

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 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 sterographic feature of

[45] `seamlessdisplay.com`,`www.wrightline.com/productDetail.asp?ProductID=14&ProductCategoryID=9&SubCategoryID=0`

LiveGraphics3D,[46] 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,"[47] the `over/under` option arranges a stereo pair vertically (with the movie for the right eye above that for the left).

A pair can be "free-view"ed 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 like the Berezin ScreenScope,[48] 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 like iO Display Systems' i-glasses[49] 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 (like that sold by Reel 3-D Enterprises[50]) 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"[51] which features such a system. A computer with

[46] `www.vis.uni-stuttgart.de/~kraus/LiveGraphics3D`
[47] `stereoscopy.com/library/waack-ch-5.html`
[48] `berezin.com/3d/screenscope.htm`
[49] `i-glassesstore.com/hmds.html`
[50] `stereoscopy.com/reel3d/screens.html`
[51] `www.ubic-u-aizu.pref.fukushima.jp/pub/system.file/system.html#3D`

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 $\mathrm{VR_4U_2C}$ and Java3D-based Just Look at Yourself!, described in the next section, automatically fill these stereo frames, either explicitly ($\mathrm{VR_4U_2C}$ 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 which 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 interfaces are illustrated by 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

Figure 16 here.

Java3D,[52] 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 & 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 16(a) right and (b) 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).

Figure 17 here.

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 rotation-invariantly, 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

[52] `java.sun.com/products/java-media/3D`

descendent geometry. As diagramed 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-" or "Gods'-eye") and elevation (side). Putting J3D `ViewPlatform`s 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 $\mathrm{VR_4U_2C}$, including all the 2.5D manipulations described in § 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 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 et al., 2004).

5.2. Dynamically Texture-Mapped Cylindrical "Goggles"

The transparency of the CG 3D space and also the photo-realistic 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 interoccular 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 ego- and exocentric perspectives. The "VR_4U_2C" multi-monitor and -display 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- and dolly-enabled non-anaglyphic stereographic QuickTime VR (SQTVR). Of course such IBR techniques will leverage advances in computational photography (Bimber, 2006), like 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 $\mathrm{VR_4U_2C}$ client runs on Macs (via the “Java Advanced Imaging” package[53]) and Windows PCs (there is no QuickTime for Java on Solaris or Linux [yet?]) 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[54] provided the x-rays described in § 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 $^{\mathrm{S}}_{\mathrm{c}}\mathrm{ha_ir^e}$ Internet Chair. In Rauterberg, M., editor, *Proc. ICEC: Third* **I***nt.* **C***onf.*

[53] apple.com/downloads/macosx/apple/java3dandjavaadvancedimagingupdate.html
[54] watanabekai.org

on **E***ntertainment* **C***omputing*, volume 3166 of *Lecture Notes in Computer Science*, pages 35–45, Eindhoven, Netherlands. Springer. ISBN 3-540-22947-7.

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 Kishino, F., Kitamura, Y., Kato, H., & Nagata, N., editors, *Proc. ICEC:* **I***nt.* **C***onf. on* **E***ntertainment* **C***omputing*, pages 511–514, Kobe Sanda, Japan. LNCS 3711; ISSN 0302-9473; ISBN 3-540-29034-6.

Adamson, C. (2005). *QuickTime for Java: a Developer's Notebook*. O'Reilly & Associates. ISBN 0-596-00822-8.

Aliaga, D. G., Funkhouser, T., Yanovsky, D., & Carlbom, I. (2003a). 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. (2003b). Interactive image-based rendering using feature globalization. In *SI3D '03: Proc. of the 2003* **S***ymp. on* **I***nteractive* **3D** *Graphics*, pages 163–170, Monterey, California. ACM Press.

Amano, K., Matsushita, F., Yanagawa, H., Cohen, M., Herder, J., Martens, W., Koba, Y., & Tohyama, M. (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. *TVRSJ:* **T***rans.* **V***irtual* **R***eality* **S***ociety of* **J***apan*, 3(1):1–12. ISSN 1344-011X; `www.u-aizu.ac.jp/~mcohen/welcome/publications/PSFC.ps`.

Arijon, D. (1991). *Grammar of the Film Language*. Silman-James Press. ISBN 1-87950-07-X.

ASCII Editing Group (2001). *iMode Java Programming*. ASCII, Tokyo. ISBN 4-7561-3727-X.

Barrilleaux, J. (2001). *3D User Interfaces with Java 3D*. Manning Publications. ISBN 1-88477-790-2.

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

Benosman, R. & Kang, S. B. (2001). *Panoramic Vision: Sensors, Theory, and Applications*. Springer. ISBN 0-387-95111-3.

Bimber, O. (2006). *(IEEE) Computer: Special Issue on Computational Photography*, 39(8).

Bolhassan, A., Cohen, M., & Martens, W. L. (2004). $\mathrm{VR_4U_2C}$: A Multiuser Multiperspective Panorama and Turnorama Browser Using QuickTime VR and Java Featuring Multimonitor and Stereographic Display. *TVRSJ:* ***T**rans.* ***V**irtual* ***R**eality* ***S**ociety of* ***J**apan*, 9(1):69–78. ISSN 1344-011X.

Bolhassan, N. A., Cohen, M., Fernando, O. N., Kamada, T., Martens, W. L., Osaka, H., & Yoshikawa, T. (2002). "Just Look At Yourself!": Stereographic Exocentric Visualization and Emulation of Stereographic Panoramic Dollying. In *Proc. ICAT:* ***I**nt.* ***C**onf. on* ***A**rtificial Reality and* ***T**ele-Existence*, pages 146–153, Tokyo. vrsj.t.u-tokyo.ac.jp/ic-at/ICAT2003/papers/02146.pdf.

Chan, Y.-F., Fok, M.-H., Fu, C.-W., Heng, P.-A., & Wong, T.-T. (1999). A Panoramic-Based Walkthrough System Using Real Photos. In *PG'99: Proc.* 7^{th} ***P**acific Conf. on Computer* ***G**raphics and Applications*, page 231, Korea. doi.ieeecomputersociety.org/10.1109/PCCGA.1999.803367.

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

Chen, S. E. & Williams, L. (1993). View interpolation for image synthesis. In *Computer Graphics (Proc.* SIGGRAPH*'93)*, pages 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. In *VRST '98: Proc. of ACM Symp. on Virtual Reality Software and Technology*, pages 81–90, Taipei, Taiwan. ACM Press. ISBN 1-58113-019-8.

Cohen, M. (1994). Cybertokyo: a Survey of Public VRtractions. *Presence: Teleoperators and Virtual Environments*, 3(1):87–93. ISSN 1054-7460.

Cohen, M. (1998). Quantity of presence: Beyond person, number, and pronouns. In Kunii, T. L. & Luciani, A., editors, *Cyberworlds*, chapter 19, pages 289–308. Springer-Verlag, Tokyo. ISBN 4-431-70207-5; `www.u-aizu.ac.jp/~mcohen/welcome/publications/QuantityOfPresence.pdf`.

Cohen, M. (2003). The Internet Chair. *IJHCI:* **I***nt.* **J***. of* **H***uman*-**C***omputer* **I***nteraction (Special Issue: Mediated Reality)*, 15(2):297–311. `www.u-aizu.ac.jp/~mcohen/welcome/publications/ic4.ps`; ISSN 1044-7318.

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? In *Proc. ICAT:* **I***nt.* **C***onf. on* **A***rtificial Reality and* **T***ele-Existence*, pages 162–169, Tokyo. `vrsj.t.u-tokyo.ac.jp/ic-at/ICAT2003/papers/02162.pdf`.

Cohen, M. & Sasa, K. (2000). An interface for a **s**oundscape-**s**tabilized **s**piral-**s**pring **s**wivel-**s**eat. In Kuwano, S. & Kato, T., editors, *Proc. WESTPRAC VII:* 7th **West***ern* **P***acific* **R***egional* **A***coustics* **C***onf.*, pages 321–324, Kumamoto, Japan. ISBN 4-9980886-1-0 and 4-9980886-3-7.

Cortázar, J. & Antonioni, M. (1966). Blowup. Metro-Goldwyn-Mayer.

Cruz-Neira, C., Sandin, D., & DeFanti, T. (1993). Surround-Screen Projection-Based Virtual Reality: Design and Implementation of the CAVE. In *Computer Graphics (Proc.* SIGGRAPH*'93)*, pages 135–142.

Darsa, L., Silva, B. C., & Varshney, A. (1997). Navigating Static Environments using Image-Space Simplification and Morphing. In *Proc. of Symp. on Interactive 3D Graphics*, pages 25–34, Providence, RI. ISBN 0-89791-884-3.

Davis, E. T. & Hodges, L. F. (1995). Human stereopsis, fusion, and stereoscopic virtual environments. In Barfield, W. & Furness III, T. A., editors, *Virtual Environments and Advanced Interface Design*, pages 145–174. Oxford University Press. ISBN 0-19-507555-2.

Dick, P. K. & Scott, R. (1982). Blade Runner. Warner Bros.

Diner, D. B. & Fender, D. H. (1993). *Human Engineering in Stereoscopic Viewing Devices*. Plenum Press, New York.

Drouin, F. (1995). *The Stereoscope and Stereoscopic Photography*. Reel 3-D Enterprises. ISBN 0-939617-02-1.

Duminduwardena, U. C. & Cohen, M. (2004). Controlling the $^{S}_{c}ha_{i}r^{e}$ Internet Chair with a Mobile Device. In *Proc. CIT: Fourth Int. Conf. on* **C***omputer* **I***nformation* **T***echnology*, pages 215–220, Wuhan, China.

Favalora, G. E. (2005). Volumetric 3D Displays and Application Infrastructure. *(*IEEE*) Computer*, 38(8):37–45.

Feng, Y. & Zhu, J. (2001). *Wireless Java Programming with J2ME*. Sams. ISBN 0-67232135-1.

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. In *Proc. MNSA: Int. Wkshp. on* **M***ultimedia* **N***etwork* **S***ystems and* **A***pplications (in conjunction with ICDCS: 24th* **I***nt.* **C***onf. on* **D***istributed* **C***omputing* **S***ystems)*, pages 170–175, Hachiōji, Tokyo. ISBN 0-7695-2087-1.

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 *Trans. on Information and Systems*, E89-D(1):73–87. `ietisy.oxfordjournals.org/cgi/content/refs/E89-D/1/73`.

Fernando, O. N. N., Cohen, M., Dumindawardana, U. C., & Kawaguchi, M. (2007). Duplex narrowcasting operations for multipresent groupware avatars on mobile devices. *IJWMC:* **I***nt.* **J***. of* **W***ireless and* **M***obile* **C***omputing (Special Issue on Mobile Multimedia Systems and Applications)*, 1(5). ISSN 1741-1084, 1741-1092.

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. In *Proc. ICIA'05:* **I***nt.* **C***onf. on* **I***nformation and* **A***utomation*, Colombo, Sri Lanka. `www.ent.mrt.ac.lk/iml/ICIA2005/Papers/JP001CRC.pdf`.

Fu, C.-W. & Heng, P.-A. (1998). A Fast Algorithm to Achieve Perceptually Acceptable Views during Multi–Node Panoramic Walk–through. In *Workshop on Computer Graphics*, pages 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. In *Proc. of the* 10^{th} *Eurographics Workshop on Rendering*, pages 169–182, Granada, Spain.

Fukuoka, T., editor (2004). *Beginning J2ME: From Novice to Professional*. Ascii. ISBN 4-7561-4503-5.

Gordon, R. & Talley, S. (1999). *Essential JMF— Java Media Framework*. Prentice Hall. ISBN 0-13-080104-6.

Hayashi, R., Nakaizumi, F., Yano, H., & Iwata, H. (2005). Development of a full-surround stereo spherical immersive display using multiple projectors. *TVRSJ:* **T***rans.* **V***irtual* **R***eality* **S***ociety of* **J***apan*, 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. In *Proc.* 4^{th} *IEEE Virtual Reality Annual Int. Symp. (VRAIS'98)*, pages 234–241, Atlanta, GA.

Hirouchi, T. (2004). *Java 3D Graphics*. Serendip, Tokyo. ISBN 4-7978-2055-1 (in Japanese).

Hoeben, A. & JanStapper, P. (2006). Taking clues from the world outside: navigating interactive panoramas. *PUC:* **P***ersonal and* **U***biquitous* **C***omputing*, 10(2–3):122–127. ISSN 1054-7460.

Horstmann, C. S. & Cornell, G. (2001). *Core Java 2, Volume I — Fundamentals*. Prentice Hall. ISBN 0-13-089468-0.

Huang, H.-C. & Hung, Y.-P. (1998). Panoramic Stereo Imaging System with Automatic Disparity Warping and Seaming. *Graph. Models Image Process.*, 60(3):196–208. ISSN 1077-3169.

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. In *Proc. CIT: Fifth Int. Conf. on* **C***omputer and* **I***nformation* **T***echnology*, pages 553–557, Shanghai. ISBN 0-7695-2432-X.

Jacobs, C. (2004). *Interactive Panoramas: Techniques For Digital Panoramic Photography*. Springer-Verlag. ISBN 3-540-21140-3.

Jorke, H. & Fritz, M. (2003). A New Stereoscopic Visualization Tool by Wavelength Multiplex Imaging. In *Proc. Electronic Displays*, Wiesbaden.

Kanno, K., Fernando, N., Bolhassan, A., Narita, S., & Cohen, M. (2006). "Personal Practically Panoramic" Multimodal Interfaces. In Simon, A., Welch, G., & Bolas, M., editors, *EDT: Proc. IEEE VR'06 Workshop on* **E***merging* **D***isplay* **T***echnologies*, pages 37–40, Alexandria, Virginia; USA. `sonic.u-aizu.ac.jp/spatial-media/Videos/keitai+Schaire2.mov`.

Kanno, T., Cohen, M., Nagashima, Y., & Hoshino, T. (2001). Mobile control of multimodal groupware in a distributed virtual environment. In Tachi, S., Hirose, M., Nakatsu, R., & Takemura, H., editors, *Proc. ICAT:* **I***nt.* **C***onf. on* **A***rtificial Reality and* **T***ele-Existence*, pages 147–154, Tokyo: University of Tokyo. ISSN 1345-1278; `sklab-www.pi.titech.ac.jp/~hase/ICATPHP/upload/39_camera.pdf`; `vrsj.t.u-tokyo.ac.jp/ic-at/ICAT2003/papers/01147.pdf`.

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

Kawakita, Y., Hamaguchi, Y., Tsukamoto, A., & Miyazaki, T. (2000). Generation of panoramic stereo images from movie using single video camera. *J. 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*. Peachpit Press, Berkeley, CA. ISBN 0-201-69684-3.

Knudsen, J. (2003). *Wireless Java: Developing with J2ME*. Apress, second edition. ISBN 1-59059-077-5.

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

Kontio, M. (2003). *Mobile Java with J2ME*. IT Press. ISBN 951-826-554-2.

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. In *ISMR:* ***I****nt.* ***S****ymp. on* ***M****ixed* ***R****eality*, pages 205–6, Tokyo.

Lea, D. (2000). *Concurrent Programming in Java*. Addison-Wesley, second edition. ISBN 0-201-31009-0.

Li, S. & Knudsen, J. (2005). *Beginning J2ME: From Novice to Professional*. Apress, third edition. ISBN 1-159059479-7.

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

Maeda, H., Tanikawa, T., Yamashita, J., Hirota, K., & Hirose, M. (2004). Real World Video Avatar: Transmission and Presentation of Human Figure. In *Proc. IEEE Virtual Reality Conf.*, pages 237–238, Chicago, IL.

Mahmoud, Q. H. (2002). *Wireless Java*. O'Reilly & Associates. ISBN 0-596-00243-2.

Maremaa, T. & Stewart, W. (2005). *QuickTime for Java: a Programmer's Guide to Building Multimedia Applications with Java*. Morgan Kaufmann, 2nd edition. ISBN 155860927X.

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

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

McAllister, D. F., editor (1993). *Stereo Computer Graphics and Other True 3D Technologies*. Princeton University Press. ISBN 0-691-08741-5.

McMillan, L. & Bishop, G. (1995a). Plenoptic modeling: an image-based rendering system. In *SIGGRAPH'95: Proc. of the 22nd Annual Conf. on Computer graphics and interactive techniques*, pages 39–46, New York. ACM Press.

McMillan, L. & Bishop, G. (1995b). Plenoptic Modeling: An Image-Based Rendering System. *Computer Graphics*, 29(Annual Conference Series):39–46.

Mikuriya, T., Shimizu, M., & Cohen, M. (2001). A collaborative virtual environment featuring multimodal information controlled by a dynamic map. *3D Forum: J. of Three Dimensional Images*, 15(1):133–136. ISSN 1342-2189.

Milgram, P. & Coquhoun, Jr., H. (1999). A Taxonomy of Real and Virtual World Display Integration. In Ohta, Y. & Tamura, H., editors, *Mixed Reality: Merging Real and Virtual Worlds*, chapter One, pages 5–30. Omsha, Springer-Verlag. ISBN 4-274-90283-8 and 3-540-65623-5.

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

Mohr, M. (2004). Spatial Narrative 'The Island of Misfit Toys'. In *Computer Graphics (Proc.* SIGGRAPH*'04)*, Los Angeles. `www.matthewmohr.com/spa_description.php`.

Naemura, T., Kaneko, M., & Harashima, H. (1998). Multi-User Immersive Stereo in Projection-Based Virtual Reality System. In *Proc. 3rd VRSJ Annual Conf.*, pages 57–60.

Nagai, T., Cohen, M., Murakami, Y., & Moriguchi, Y. (2006). Networked driving simulator with force display, rotating motion platform, and dual steering. In *Proc. HC2006: Ninth Int. Conf. on* **H***uman and* **C***omputer*, pages 25–30, Aizu-Wakamatsu, Japan.

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

Naimark, M. (1997). A 3D Moviemap and a 3D Panorama. In *Proc. SPIE/SPSE Electronic Imaging Conf.*, San Jose, CA. Vol. 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. In *Proc.* IEEE *Computer Vision and Pattern Recognition (CVPR)*. www.cs.columbia.edu/CAVE/publinks/nayar_cvpr_1997.pdf.

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

Ota, A. (2003). *Yasasiku Wakaru Java 3D*. Software Technology. Gijutsu Hyoron Sya. ISBN 4-7741-1669-6 (in Japanese).

Palmer, I. (2001). *Essential Java 3D fast: developing 3D graphics applications*. Springer. ISBN 1-85233-394-4.

Pausch, R., Burnette, T., Brockway, D., & Weiblen, M. E. (1995). Locomotion in virtual worlds via flight into hand-held miniatures. In *Computer Graphics (Proc.* SIGGRAPH*'95)*, pages 399–400, Los Angeles. www.cs.cmu.edu/~stage3/publications/95/conferences/siggraph/paper.html, www.cs.cmu.edu/~stage3/publications/95/conferences/chi/paper.html.

Peleg, S. & Ben-Ezra, M. (1999). Stereo panorama with a single camera. In *Proc. CVPR: (*IEEE *Computer Society Conf. on)* **C***omputer* **V***ision and* **P***attern* **R***ecognition*, pages 395–401, Ft. Collins, Colorado; USA.

Peleg, S., Ben-Ezra, M., & Pritch, Y. (2001). Omnistereo: Panoramic Stereo Imaging. *(*IEEE*) Trans. 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). In *OMNIVIS00: Proc. IEEE Wkshp. on Omnidirectional Vision*, pages 54–61, Hilton Head Island, SC; USA. ISBN 0-7695-0704-2.

Rademacher, P. & Bishop, G. (1998). Multiple-Center-of-Projection Images. In *Computer Graphics (Proc.* SIGGRAPH*'98)*, pages 199–206, Orlando, Florida.

Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S., & Forlines, C. (2003). iLamps: geometrically aware and self-configuring projectors. *ACM Trans. Graph.*, 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. ISSN 1054-7460.

Salvendy, G., editor (1997). *Handbook of Human Factors and Ergonomics*. John Wiley & Sons, second edition. ISBN 0-471-11690-4.

Sasaki, M. & Cohen, M. (2004). Dancing Music: Integrated MIDI-Driven Synthesis and Spatialization for Virtual Reality. In *AES: Proc.* **A***udio* **E***ngineering* **S***ociety Conv.*, San Francisco. Preprint 6316 (R-3).

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

Selman, D. (2002). *Java 3D Programming*. Manning. ISBN 1-930110-35-9.

Shade, J. W., Gortler, S. J., He, L.-W., & Szeliski, R. (1998). Layered Depth Images. In *Computer Graphics (Proc.* SIGGRAPH*'98)*, pages 231–242, Orlando, Florida.

Shimada, D., Tanahashi, H., Kato, K., & Yamamoto, K. (2001). Extract and display moving object in all direction by using stereo omnidirectional system (sos). In *3DIM: Proc. Third Int. Conf. on 3-D Digital Imaging and Modeling*, pages 42–47.

Shum, H.-Y., Kalai, A., & Seitz, S. M. (1999). Omnivergent stereo. In *Seventh Int. Conf. on Computer Vision*, pages 22–29. ISBN 0-7695-0164-8.

Snavely, N., Seitz, S. M., & Szeliski, R. (2006). Photo tourism: Exploring photo collections in 3d. In *Trans. on Graphics (Proc.* SIGGRAPH*'06)*, volume 25, pages 835–846. http://phototour.cs.washington.edu/Photo_Tourism.pdf.

Sowizral, H., Rushforth, K., & Deering, M. (1998). *The Java 3D API Specification (Japanese edition)*. Addison-Wesley. ISBN 4-7561-3017-8.

Sowizral, H., Rushforth, K., & Deering, M. (2000). *The Java 3D API Specification*. Addison-Wesley, second edition. ISBN 0-201-71041-2.

Stanney, K. M., editor (2002). *Handbook of Virtual Environments: Design, Implementation, and Applications*. Human Factors and Ergonomics. Lawrence Erlbaum Associates, Mahway, NJ. ISBN 0-8058-3270-X.

Stoakley, R., Conway, M. J., & Pausch, R. (1995). Virtual reality on a wim: interactive worlds in miniature. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, pages 265–272, Denver. ACM Press/Addison-Wesley Publishing Co.

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

Tanaka, K., Hayashi, J., Endo, T., & Tachi, S. (2003). A method for panoramic stereo image acquisition. In *Proc. ICAT: Thirteenth Int. Conf. on Artificial Reality and Telexistence*, pages 257–262, Tokyo: Keio University. ISSN 1345-1278.

Tanaka, S., Matsubara, S., Monobe, K., Miura, S., Miyawaki, Y., Hattori, H., Hiura, K., Kobayashi, R., & Yamada, K. (2002). *Java3D Graphics Introduction*. Morikita, Tokyo. ISBN 4-627-84391-7 (in Japanese).

Topley, K. (2002). *J2ME in a Nutshell*. O'Reilly & Associates. ISBN 0-596-00253-X.

Tzavidas, S. & Katsaggelos, A. K. (2005). A Multicamera Setup for Generating Stereo Panoramic Video. *IEEE Trans. on Multimedia*, 7(5):880–890.

Vacca, J. R. (2002). *I-Mode Crash Course*. McGraw-Hill. ISBN 0-07-138187-2.

Vineyard, J. (2000). *Setting Up Your Shots: great camera moves every filmaker should know*. Michael Wiese Productions. ISBN 0-941188-73-6.

Waldsmith, J. (2002). *Digital Image Processing*. Krause Publications, second edition. ISBN 0-87349-409-1.

Walrath, K. & Campione, M. (1999). *The JFC Swing Tutorial*. Addison-Wesley. ISBN 0-201-43321-4.

Walsh, A. E. & Gehringer, D. (2002). *Java 3D API Jump-Start*. Prentice-Hall. ISBN 0-13-034076-6.

Watkinson, J. (2001). *Convergence in Broadcast and Communications Media*. Focal Press. ISBN 0-240-51509-9.

Yamazaki, Y. (2001). *How to make an iAppli*. SoftBank, Tokyo. ISBN 4-7973-1573-3.

(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)

Yendo, T., Kawakami, N., & Tachi, S. (2005). Seelinder: the cylindrical light-field display. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Emerging technologies*, page 16, New York. ACM Press. `laputanlogic.com/articles/2004/11/028-0001-8703.html`.

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

Image Quality

live action movie

digital movie

QTVR

IBR (3D from 2D image interpolation)

texture-mapped CG

conventional computer graphics

Interactivity

Figure 2. Image Quality vs. Interactivity

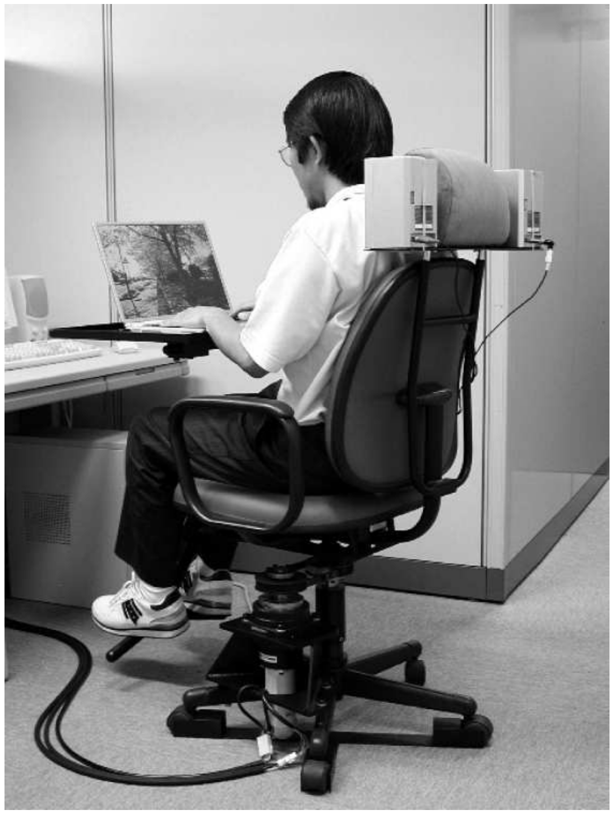
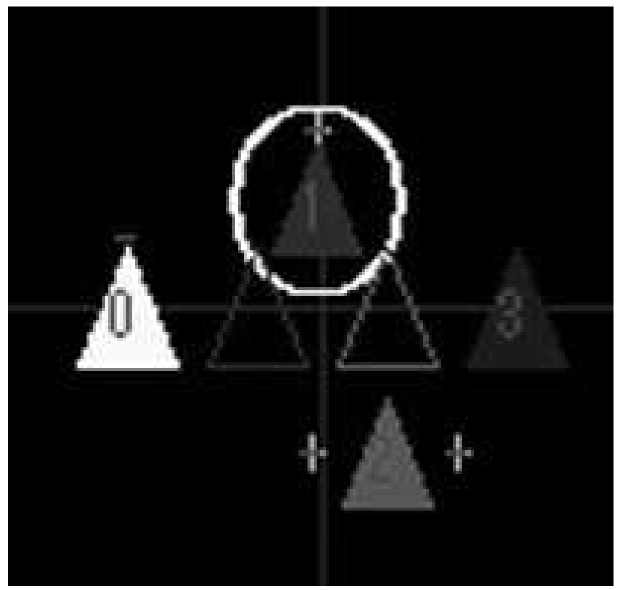


Figure 3. $^{\mathrm{S}}_{\mathrm{c}}\mathrm{ha_{i}r^{e}}$ Internet Chair integrated with $\mathrm{VR_4U_2C}$: The swivel chair is a rotary motion platform. The panorama displayed by the laptop computer (with wireless internet connectivity) is propriocentrically anchored, panning opposite to the chair 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.[55])

[55] `www.mechtec.co.jp`

(a) Egocentric Panoramic Browser “$\mu\mathrm{VR_4U_2C}$”: 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 “$\imath$·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 Yoshie Tanno.)

Figure 4. These two mobile applications for NTT DoCoMo iαppli mobile phones are integrated with $\mathrm{VR_4U_2C}$ and Just Look at Yourself!, and can reflect and control the workstation displays.

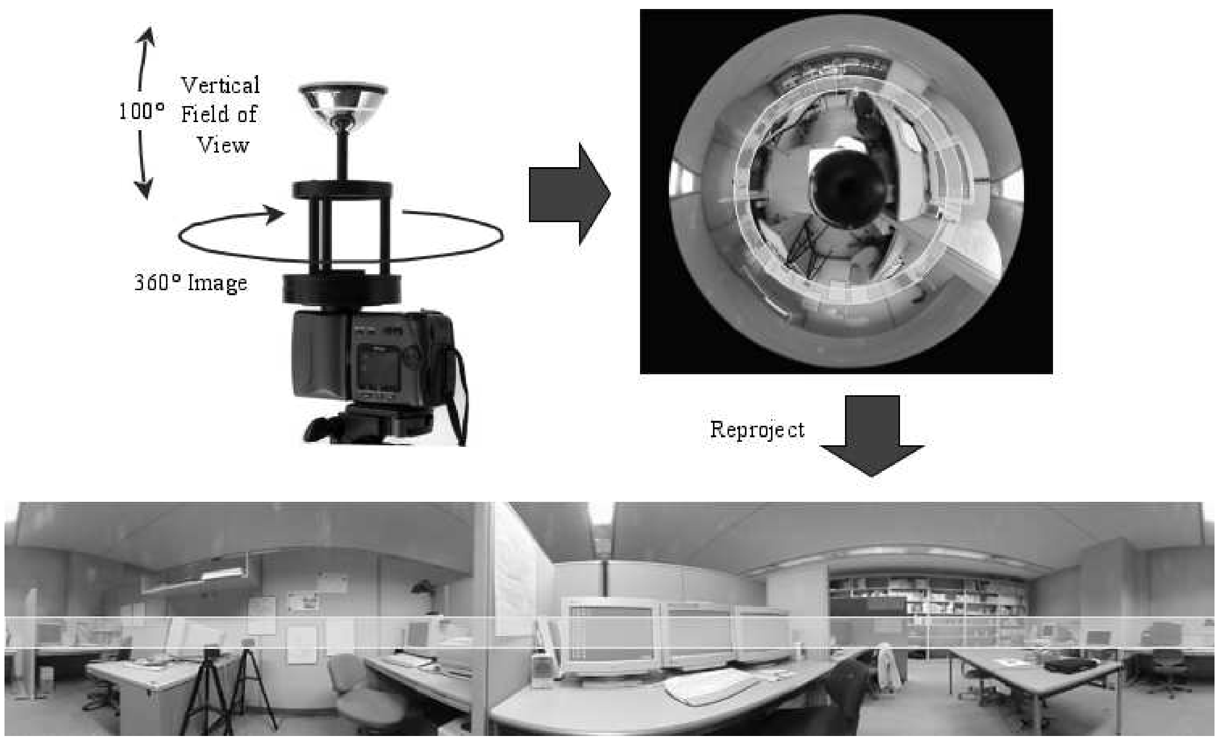
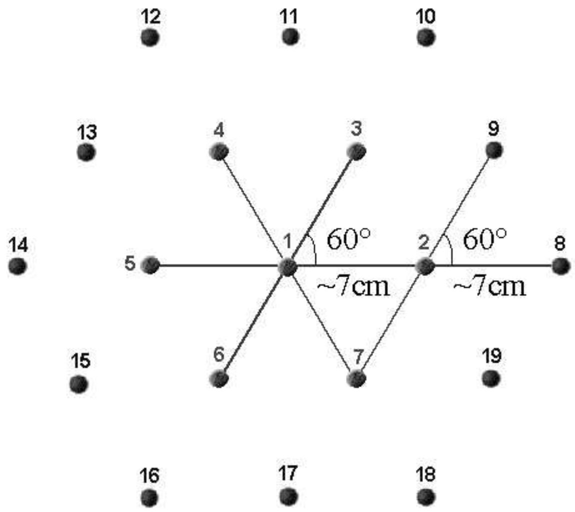
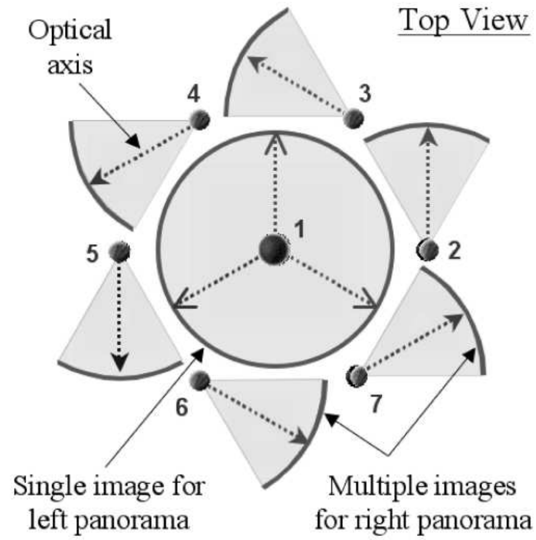


Figure 5. Preparing a Panoramic Scene. (Equatorial bands are superimposed to show corresponding horizons in the various projections.)



(a) Capture Lattice



(b) Node Selection

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.

Figure 7. Example of Left/Right Stereo Images. Note binocular parallax viewpoint displacement, causing offset of objects like the paper on the left edge.

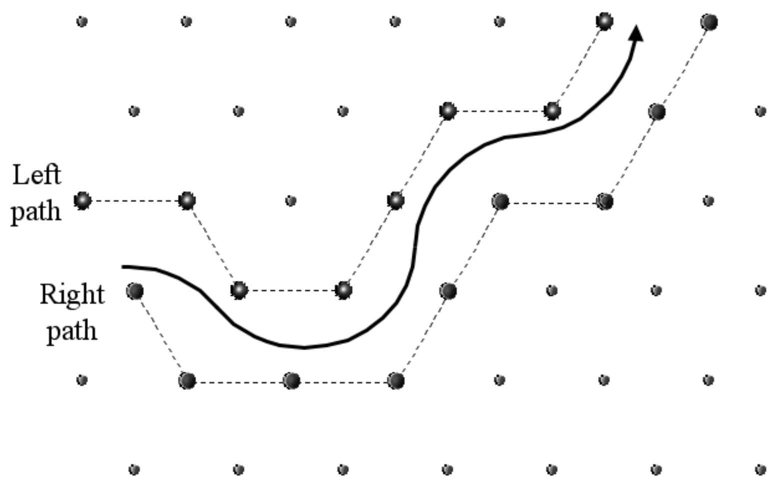


Figure 8. Left/Right Stereo Paths Determined from Walk-through Meander

Figure 9. Equipment for Capturing Turnoramas

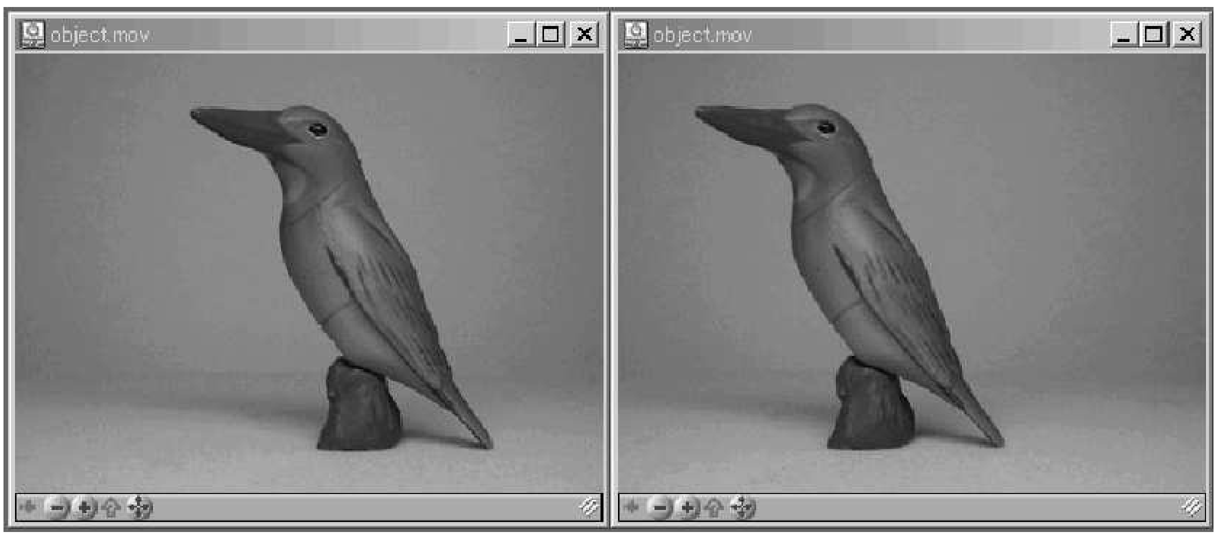


Figure 10. Example of Left/Right Stereo Turnoramic Images

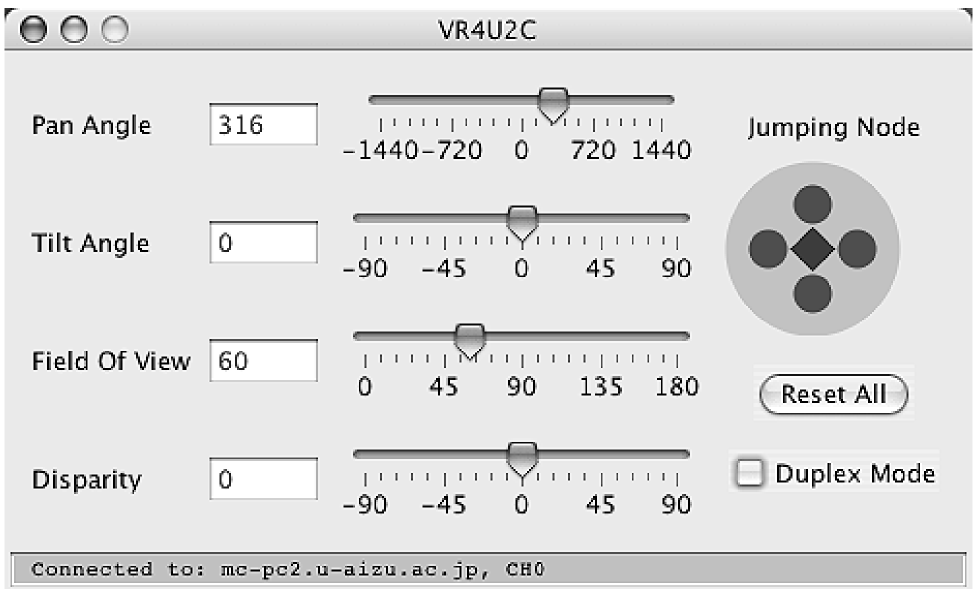
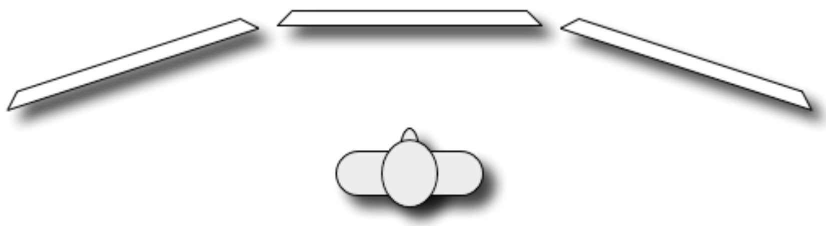


Figure 11. $\mathrm{VR_4U_2C}$ Navigator with Continuous & Numeric Controllers including Dolly & Track Rosette

Figure 12. Multiple Synchronized Windows

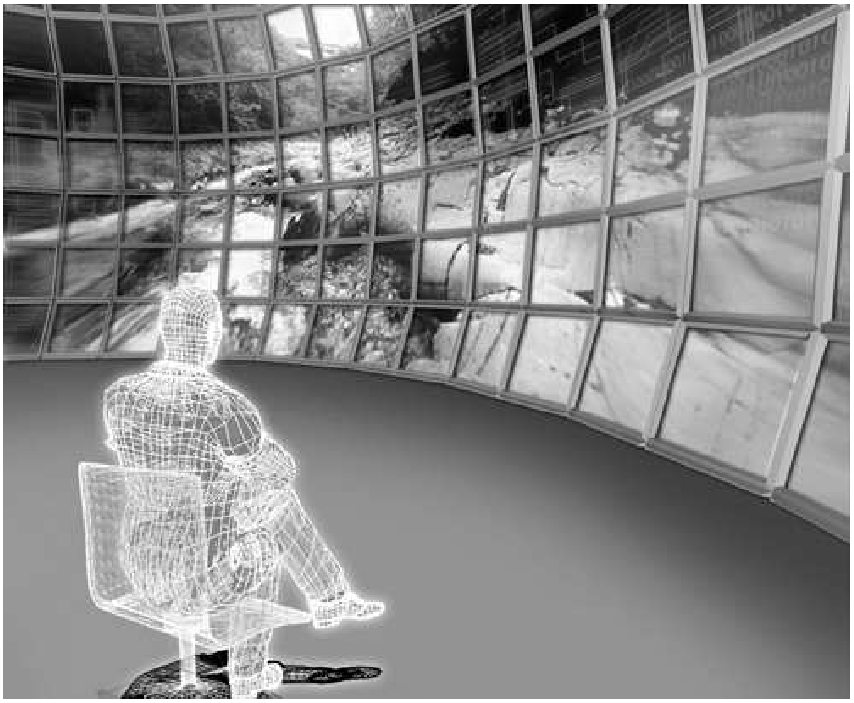


(a) Schematic

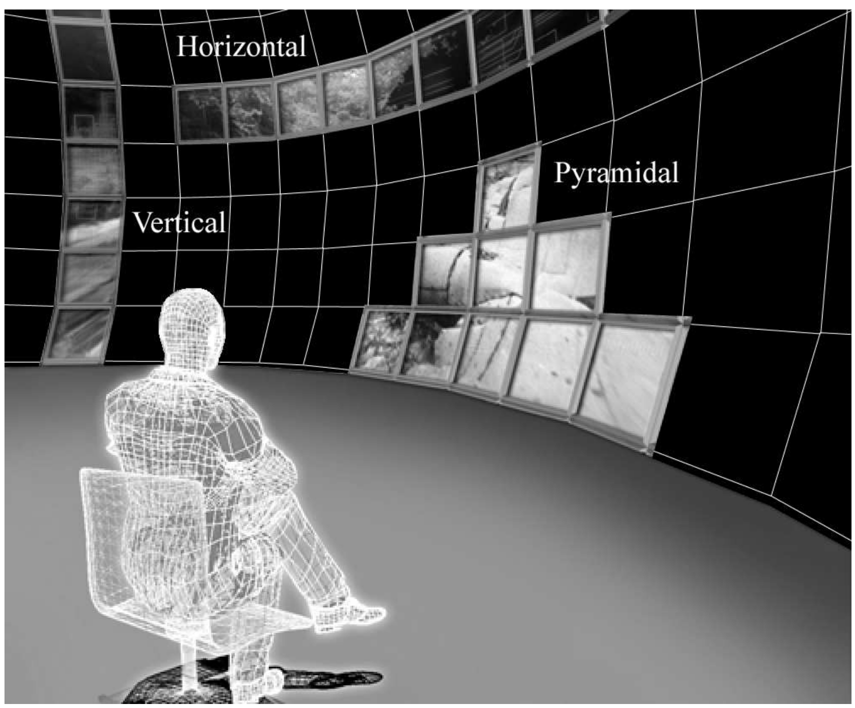
(b) Typical View

Figure 13. Multiple Synchronized Monitors: Panoramic Panorama

[56] aizu.com

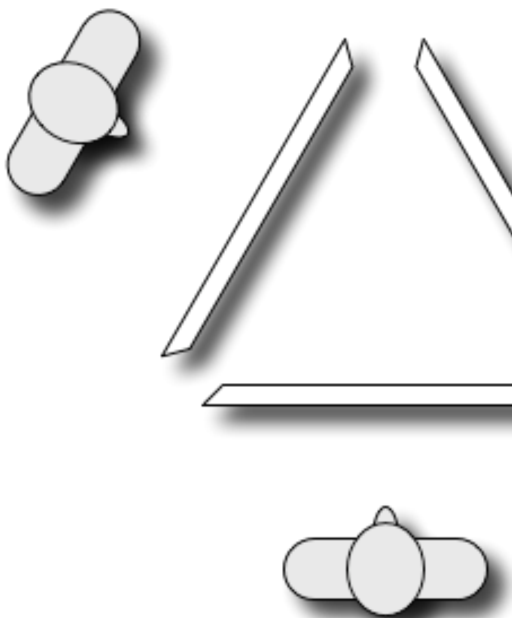


(a) Immersive Imagery



(b) Tiling Styles

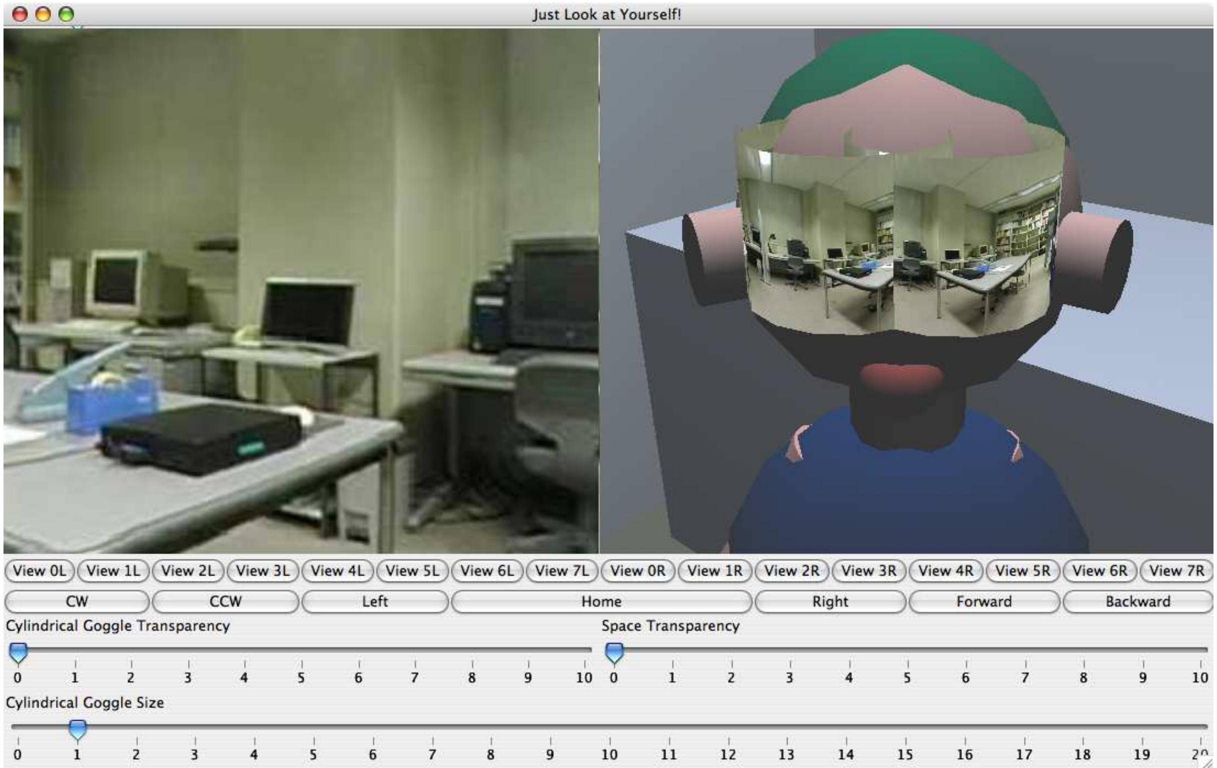
Figure 14. Graphic Depiction of a Curved "Video-Wall" Configuration of Visual Display Devices (Graphics by "Eyes, Japan."[56])



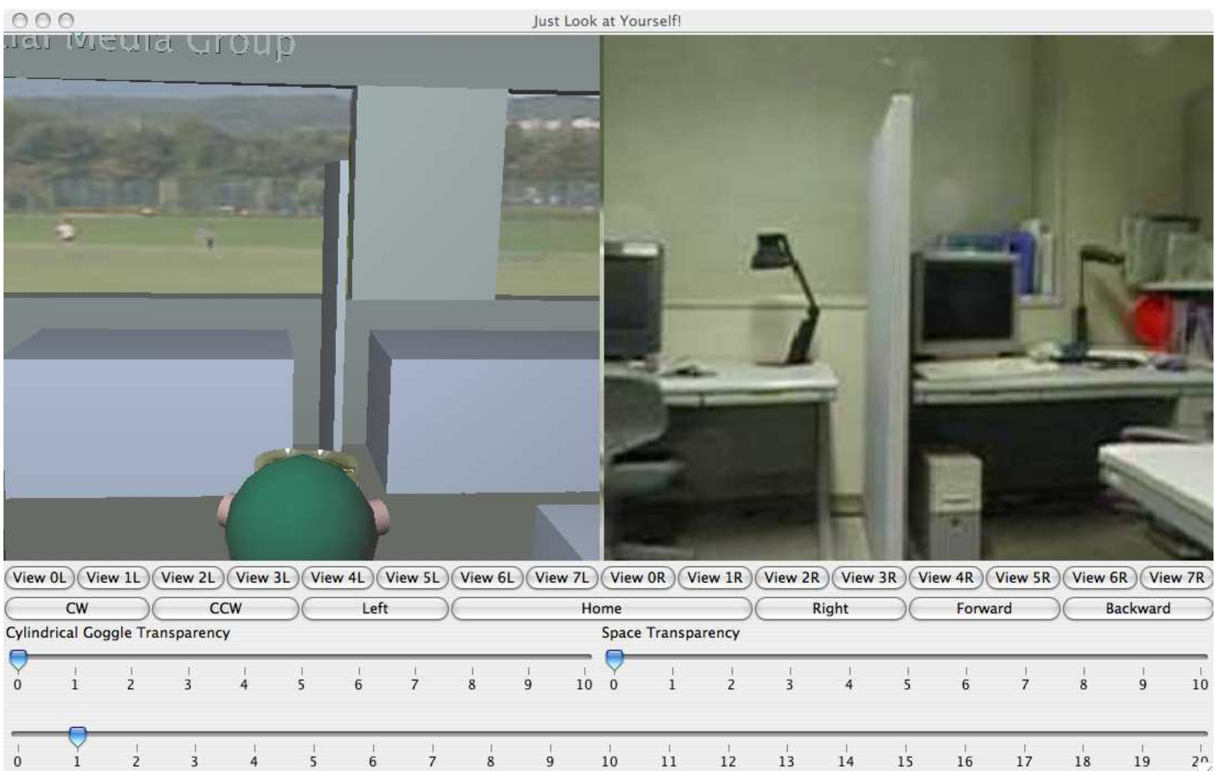
(a) Schematic

(b) Juxtaposed Views

Figure 15. Multiple Synchronized Monitors: Turnoramic Turnorama



(a) Both panes are configurable to show an endo-, ego-, or exocentric stereographic IBR, stereographic 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

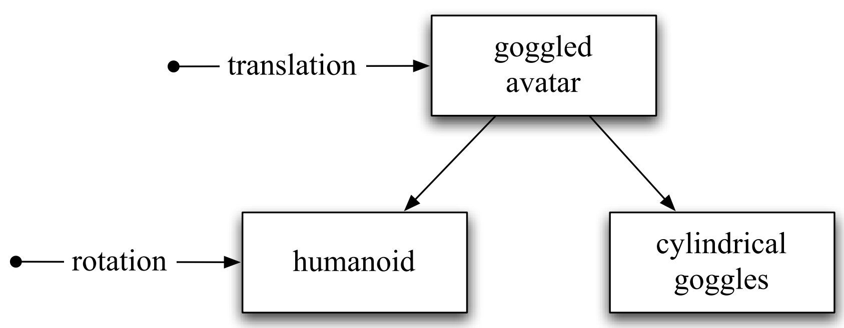


Figure 17. Repositioning Routing by Scene Graph Branch

Contents

List of Figures

List of Tables