Personal Varrier: Autostereoscopic Virtual Reality Display for Distributed Scientific Visualization

Tom Peterka[1], Daniel J. Sandin[1], Jinghua Ge[1], Javier Girado[1], Robert Kooima[1], Jason Leigh[1], Andrew Johnson[1], Marcus Thiebaux[2], Thomas A. DeFanti[1]

[1] Electronic Visualization Laboratory
University of Illinois at Chicago

[2] Information Sciences Institute
University of Southern California

Address correspondence to:
Tom Peterka
Electronic Visualization Laboratory
University of Illinois at Chicago
Room 2032 Engineering Research Facility (ERF)
842 W. Taylor Street, Chicago, IL 60607
312-996-3002
tpeterka@evl.uic.edu

Abstract

As scientific data sets increase in size, dimensionality, and complexity, new high resolution, interactive, collaborative networked display systems are required to view them in real-time. Increasingly, the principles of virtual reality (VR) are being applied to modern scientific visualization. One of the tenets of VR is stereoscopic (stereo or 3d) display, however the need to wear stereo glasses or other gear to experience the virtual world is encumbering and hinders other positive aspects of VR such as collaboration. Autostereoscopic (autostereo) displays present imagery in 3d without the need to wear glasses or other gear, but few qualify as VR displays. The Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago (UIC) has designed and built a single-screen version of its 35-panel tiled Varrier display, called Personal Varrier. Based on a static parallax barrier and the Varrier computational method, Personal Varrier provides a quality 3d autostereo experience in an economical, compact form factor. The system debuted at iGrid 2005 in San Diego, CA, accompanied by a suite of distributed and local scientific visualization and 3d teleconferencing applications. The CAVEwave National LambdaRail (NLR) network was vital to the success of the stereo teleconferencing.

Keywords: Varrier, Personal Varrier, virtual reality, autostereoscopy, autostereo, 3D display, camera-based tracking, parallax barrier

1. Introduction

First person perspective, 3d stereoscopic vision, and real-time interactivity all contribute to produce a sense of immersion in VR displays. These features can afford greater understanding of data content when VR is applied to the field of scientific visualization, particularly when the data sets are large, complex, and highly dimensional. As scientific data increases in size, computation and display have been de-coupled through mechanisms such as amplified collaboration environments [13] which often contain stereo display devices.

However, active and passive stereo displays require stereo glasses to be worn, which inhibit the collaborative nature of modern research. Face to face conversation and 3d teleconferencing are impoverished because the 3d glasses or headgear block lines of sight between participants. Even in solitary setttings, shutter or polarized glasses can be uncomfortable when worn for long periods of time, block peripheral vision, and interfere with non-stereo tasks such as typing or reading.

Autostereo displays seek to address these issues by providing stereo vision without the need for special glasses or other gear. In recent years, autostereo displays have gained popularity in the data display market. EVL first introduced the Cylindrical Varrier system at IEEE VR in 2004, and published its results [11] earlier this year. The 35-panel Cylindrical Varrier is a high-end clustered display system that delivers an immersive experience, but its size, complexity, and cost make it difficult and expensive to transport to conferences and to deploy in the locations of EVL's collaborating partners. To this end, EVL developed a desktop Varrier form factor called Personal Varrier, which debuted at iGrid 2005 in San Diego, CA.

Personal Varrier features a camera-based face recognition and tracking system, autostereo teleconferencing capability, a modest size desktop form factor, 3d wand interactivity, support for distributed computing applications, and 3d animation capability. These features combined with improved 3d image quality and lack of stereo glasses make Personal Varrier a VR

display that can be used comfortably for extended periods of time. The compact footprint and relatively modest cost make it a system that is easier to deploy in terms of space and budget than the Cylindrical Varrier. Yet, as was seen at iGrid, Personal Varrier provides an interactive, engaging VR experience.

A key component to the success of the networked applications shown at iGrid 2005 is the National LambdaRail (NLR). NLR is a greater than 10,000-mile owned lit-fiber infrastructure for and by the US research and education community, capable of providing up to forty 10Gb/s wavelengths. EVL in Chicago is now 10 Gigabit Ethernet-connected to UCSD via Seattle on its own private wavelength on the NLR. Called the CAVEwave, because CAVE[TM] royalty money has funded it, this link is dedicated to research. The CAVEwave is available to transport experimental traffic between federal agencies, international research centers, and corporate research projects that bring gigabit wavelengths to Chicago, Seattle, and San Diego.

2. Background

Autostereo displays can be produced by a variety of methods, including optical, volumetric, and lenticular / barrier strip, but lenticular and barrier strip remain the most popular methods due to their ease of manufacture. A comprehensive survey of autostereo methods can be found [6] and more recently in [12].
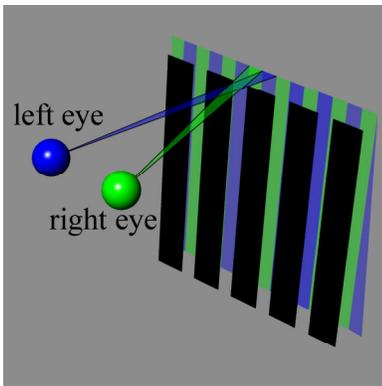
Figure 1: Varrier is based on a static parallax barrier. The left eye is represented by blue stripes and the right eye with green; these image stripes change with viewer's motion while the black barrier remains fixed.

Varrier (variable or virtual barrier) is the name of both the system and a unique computational method. The Varrier method utilizes a static parallax barrier, which is fine-pitch alternating sequence of opaque and transparent regions printed on photographic film as in Figure 1. This film is then mounted in front of the LCD display, offset from it by a small distance. The displayed image is correspondingly divided into similar regions of left and right eye perspective views, such that all of the left eye regions are visible only by the left eye and likewise for the right eye regions. The brain is thus simultaneously presented with two disparate views, which are fused into one stereo or 3d image. An alternative to the relatively simple static barrier is an active barrier built from a solid state micro-stripe display as in [7], [8]. Lenticular displays function equivalently to barrier strip displays, with the barrier strip film replaced by a sheet of cylindrical lenses as in the SynthaGram display [5].
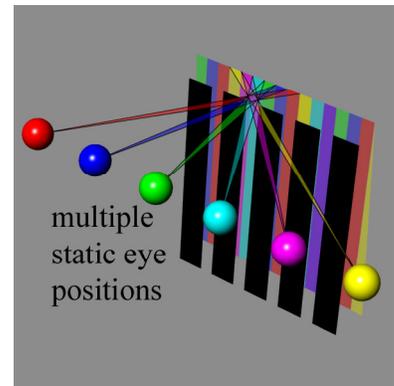
Figure 2: An alternative to Varrier's tracked 2-view method in Fig. 1 is used in some other systems: the multi-view panoramagram. Multiple views represented by multi-colored stripes are rendered from static untracked eye positions.

Parallax barrier and lenticular screens are usually mounted in a rotated orientation relative to the pixel grid to minimize or restructure the Moiré effect that results as an interference pattern between the barrier and pixel grid, [6], [16], [18].

Parallax barrier autostereo displays generally follow one of two design paradigms. Tracked systems such as Varrier [10], [11], and others such as [14], [7], [8] produce a stereo pair of views that follow the user in space, given the location of the user's head or eyes from the tracking system. They are strictly single-user systems at this time. Untracked two-view systems also exist where the user is required to maintain their position in the pre-determined "sweet spot" for which the system is calibrated. The other popular design pattern is the untracked panoramagram as in Figure 2, where a multiple sequence of perspective views is displayed from slightly varying vantage points. Multiple users can view this type of display with limited "look-around" capability. [5]. Each approach has its relative merits; a discussion of tracked and untracked autostereo can be found in [11].

In the previous barrier strip autostereo systems, interleaved images are prepared by rendering left and right eye views and then sorting the images into columns. The Varrier computational method [10], [11] interleaves left and right eye images in floating-point world space, rather than integer image space. Three different algorithms for accomplishing the spatial multiplexing are described in [11] each resulting in varying degrees of performance/quality trade-offs. In each case, a virtual parallax barrier is modeled within the scene and used together with the depth buffer to multiplex two perspectives into one resulting image. Perspective sight lines are directed from each of the eye positions, through the *virtual* barrier to the scene objects that are occluded by the virtual barrier. When the corresponding pixels are displayed and the light is filtered by the *physical* barrier, each eye receives only the light that is intended for it. With head tracking, only two views are required. Tracking latency needs to be low and frame rates need to be high to produce real-time interactivity. The result is that images are steered in space and follow the viewer. For example, in Figure 3 a green color representing the left eye image and blue for the right eye image are visible on the white card held by the user. For this experiment, a tracking sensor is held above the card and the images follow the movement of the card and sensor. In a similar way, the left and right eye images follow a user's face when tracked by the camera-based tracking system.
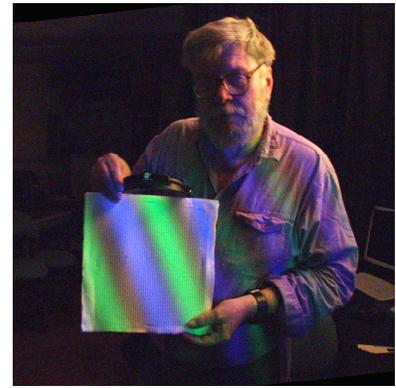
Figure 3: A test pattern consisting of green stripes for the left eye and blue for the right eye is displayed on the card held by the subject. The images are steered in space and follow the motion of the card because the card is being tracked in this experiment.

3. System configuration

Display system

The Personal Varrier LCD display device is an Apple^{TM} 30" diagonal wide-screen monitor with a base resolution of 2560 x 1600 pixels. A 22 in. x 30 in. photographic film is printed via a pre-press process containing a fine spacing of alternating black and clear lines, at a pitch of approximately 270 lines per foot. The film is then laminated to a 3/16 in. thick glass pane to provide stability, and the glass/film assembly is affixed to the front of the monitor as in Figure 4. The pitch of the barrier is near the minimum governed by the Nyquist Sampling Theorem, which dictates that at least two pixels per eye must be spanned in each line cycle. Each line cyle is approximately ¾ opaque and ¼ transparent, reducing the effective horizontal resolution to 640 pixels. The physical placement of the barrier is registered with the system entirely through software [1].

The horizontal angle of view is 50 degrees when the user is 2.5 feet from the display screen. The user is free to move in an area approximately 1.5 ft. x 1.5 ft., centered about this location. This provides a comfortable viewing distance for a seated desktop display. VR images are rendered using a computer that has dual Opteron^{TM} 64-bit 2.4 GHz processors, 4GB RAM, and 1Gbps Ethernet.

The system also includes some auxiliary equipment. Two 640x480 Unibrain^{TM} cameras are used to capture stereo teleconferencing video of the user's face. These are mounted just below the 3 cameras used for tracking. Additionally, a 6-DOF tracked wand is included, tracked by a medium-range Flock of Birds^{TM} tracker. The wand includes a joystick and 3 programmable buttons, for 3d navigation and interaction with the application. A photograph and layout of the Personal Varrier is shown in Figures 5 and 6, respectively.

Figure 4: Construction of the Varrier screen is a simple modification to a standard LCD display. The barrier strip is printed on film and laminated to a pane of glass to provide stability. The glass/film is then mounted .150 in. in front of the LCD display surface. It is this gap that provides parallax between the barrier strip and image stripes.

Camera tracking system

Tracked autostereo requires real-time knowledge of the user's 3d head position, preferably with a tether-less system that requires no additional gear to be worn for tracking. In Personal Varrier, tracking is accomplished using 3 high speed (120 frames per second, 640 x 480) cameras and several artificial neural networks (ANNs) to both recognize and detect a face in a crowded scene and track the face in real-time. More details can be found in [2], [3]. Independence from room and display illumination is accomplished using 8 infrared (IR) panels and IR band-pass filters on the camera lenses. The center camera is used for detection (face recognition) and for x,y tracking, while the outer two cameras are used for range finding (see Figure 7). 3 computers process the video data and transmit tracking coordinates to the Personal Varrier. Each computer contains dual Xeon^{TM} 3.6 GHz processors, 2GB RAM and 1 Gbps Ethernet.

Training for a new face is short and easy, requiring less than two minutes per user. The user slowly moves his or her head in a sequence or normal poses while 512 frames of video are captured which are then used to train the ANNs. User data can be stored and re-loaded, so the user needs to train the system only once. Once trained and detected, the system tracks a face at a frequency of 60 Hz. Additionally, there is approximately an 80% probability that a generic training session can be applied to other users. For example, in a conference setting such as iGrid, the system is trained for one face and the same training parameters are used for most of the participants, with overall success.

The tracking system has certain limitations. The operating range is currently restricted to 1.5 x 1.5 feet. Presently the tracking system is supervised by a trained operator, although a new interface is being prepared that will make this unnecessary. End-to-end latency is 81 ms, measured as in [4]. Most of this latency is due to communication overhead, and work is ongoing to transfer tracking data more efficiently between the tracking machines and the rendering machine.

Figure 5: A user is seated at the Personal Varrier system, designed in a desktop form factor to be used comfortably for extended periods of time. Autostereo together with camera-based tracking permits the user see 3d stereo while being completely untethered.
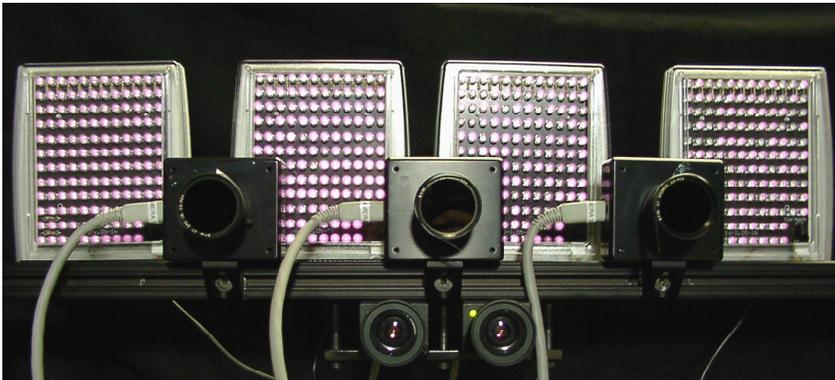
Figure 7: A close-up view is shown of 3 tracking cameras for face recognition and tracking, 4 IR illuminators to control illumination for tracking, and 2 stereo conferencing cameras for 3d autostereo teleconferencing.
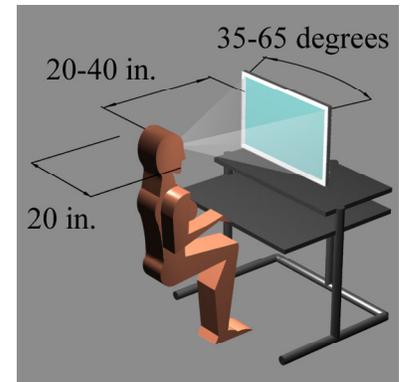
Figure 6: Personal Varrier fits a standard 36" x 30" desk. The user can move left and right 20 inches, and be a distance of between 20 to 40 inches from the display screen. Field of view is between 35 and 65 degrees, depending on the user's location.

4. Distributed and local scientific visualization demonstrations

Point sample packing and visualization engine (PPVE)
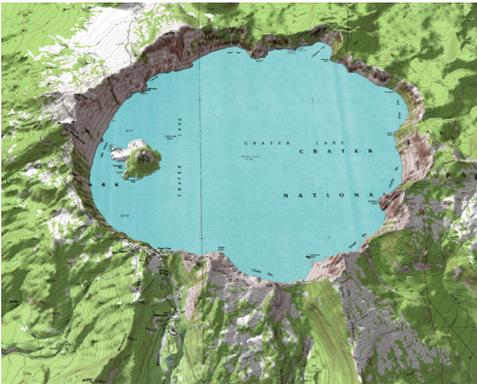
Figure 8: The PPVE application implements a client-server remote data visualization model, with the server generating texture and depth maps and the client reconstructing viewpoints by point splatting. A topography of Crater Lake is visualized above.

This application is designed to visualize highly-detailed massive data sets at interactive frame rates for the Varrier system. Since most large scientific data sets are stored on remote servers, it is inefficient to download the entire data set and store a local copy every time there is a need for visualization. Furthermore, large datasets may only be rendered at very low frame rates on a local computer, inhibiting Varrier's interactive VR experience. PPVE is designed to handle the discontinuity between local capability and required display performance, essentially de-coupling local frame rate from server frame rate, scene complexity, and network bandwidth limitations.

A large dataset is visualized in client-server mode. Personal Varrier, acting as the client sends frustum requests over network to the server, and the remote server performs on-demand rendering and sends the resulting depth map and color map back to the client. The client extracts 3d points (coordinates and color) from received depth and color maps and splats them onto the display. Points serve as the display primitive because of their simple representation, rendering efficiency, and quality [9].

Server: The server's tasks include on-demand rendering, redundancy deletion, map compression and data streaming. The server maintains a similar octree-based point representation as the client to ensure data synchronization. The server also maintains the original complete dataset, which may take one of several forms including polygonal, ray tracing, or volume rendering data. The server need not be a single machine, but can be an entire computational cluster.

Client: The client's tasks include map decompression, point extraction and decimation, point clustering and local point splatting. Additional features include incremental multi-resolution point-set growing, view-dependent level-of-detail (LOD) control, and octree-based point grouping and frustum culling. The client permits navigation of local data by the user while new data is being transmitted from the server.

A sample dataset containing 5 M triangles and a 4 K by 4 K texture map can be rendered by the server at only 1 Hz, but with PPVE, the client can reconstruct a view at 20-30 Hz. At iGrid, the transmission data rate was approximately 2 MB/s, and PPVE was run over the local network, from an on-site server. A 3d topography application was demonstrated, as shown in Figure 8.

3d video/audio teleconferencing

Scientific research today depends on communication; results are analyzed and data are visualized collaboratively, often over great distances and preferably in real-time. The multi-media collaborative nature of scientific visualization is enhanced by grid-based live teleconferencing capability as in Figure 9. In the past, this has been a 2d experience, but this changed at iGrid 2005, as conference attendees sat and discussed the benefits of autostereo teleconferencing "face to face" in real-time 3d with scientists at EVL in Chicago. Initial feedback from iGrid was positive and participants certainly appeared to enjoy the interaction. Thanks to CAVEwave 10 gigabit optical networking between UCSD and UIC, audio and stereo video were streamed in real-time, using approximately 220 Mb/s of bandwidth in each direction. End to end latency was approximately .3 s (round trip), and autostereo images of participants were rendered at 25 Hz. Raw, unformatted YUV411 encoded images were transmitted, using one half the bandwidth as compared to a corresponding RGB format. The real-time display for the most part appeared flawless even though approximately 10% of packets were dropped. The robustness is due to the high transmission and rendering rates. The sending rate was 30 Hz, while the rendering rate was approximately 25 Hz, limited mainly by pixel fill rates and CPU loading. The net result is that the occasional lost packets were negligible. The

Figure 9: A researcher at EVL chats in autostereo with a remote colleague on the Personal Varrier.

excellent network quality of service can be attributed to the limited number of hops and routers between source and destination. There were exactly 4 hosts along the route, including source and destination, so only two intermediate routers were encountered. The success of the teleconferencing is directly attributable to the 10 gigabit CAVEwave light path over the National LambdaRail, as the application simply sends raw stereo video and audio data and relies on the network to successfully deliver quality of service.

Local applications

Because network bandwidth at the iGrid conference needs to be shared among many different groups, bandwidth is scheduled and there are periods of time when local, non-distributed autostereo applications are shown. Some of these, for example the photon-beam tomography rendering, are local versions of otherwise highly parallel distributed computation and rendering. This sub-section highlights some of the local autostereo applications demonstrated at iGrid 2005 and shown in Figure 10.

The Mars rover application showcases data downloaded from the Spirit and Opportunity rover missions to Mars. Seven different panoramas, courtesy of NASA, are texture mapped in the application, selectable with a wand button. Each image is approximately 24 MB. The application also depicts a model of the Spirit rover from a third party. The application serves as a test program for Varrier, allowing the user to calibrate and control all of Varrier's optical parameters from a control panel on the console. The user can interact with the application via the keyboard and a tracked wand, permitting navigation in the direction of the wand with the joystick, and can probe data with a 3d pointing device.

The Hipparcos Explorer application displays data collected by the European Space Agency (ESA) Hipparcos satellite. The Hipparcos mission lasted from August 1989 to August 1993, although the satellite remains in Earth orbit today. During its active period it followed the Earth in its orbit around the Sun and measured the positions of 2.5 million stars to a precision of only a few milliseconds of arc. The parallax observed over the course of each year allows one to compute the distance to approximately 120 thousand stars and determine their positions in 3d space with reasonable precision. In addition, spectral type is used to produce the color of each star. The position, brightness, and color of these stars define a scene approximately 1000 light-years in diameter, which can be rendered from any viewpoint at interactive rates. Personal Varrier allows one to fly through the field of stars using the tracked wand control and discern shapes of constellations and other familiar clusters of stars, all in 3D autostereo.

Photon-beam tomography generates volumetric datasets based on sequentially angled X-ray images of small objects [17]. Using the Personal Varrier, a 3D X-ray of cricket's head was examined using a splat-based volume rendering technique. In order to generate a discernable image from a dense volume, a transfer-function is applied to each point in the volume, resulting in a color and an opacity, or brightness. Human perceptual limitations require that the percentage of volume elements that contribute to the image, or 'interest ratio', must be very small. Volume splatting leverages this feature by representing only the small percentage of cells that fall in the range of interest. At ISI, work with Scalable Grid-Based Visualization [15] applies this technique to support interactive visualization of large 4D time-series volume data. The next step will be to feed a stream of view-point independent splat sets to Varrier from a bank of remote cluster nodes, to enable fully interactive viewing of volume animations. The cricket volume is 512^3 byte scalars, and renders at 10 Hz in stereo.

Figure 10: While network bandwidth was scheduled to other iGrid demonstrations, Personal Varrier displayed local applications in autostereo. Mars Rover (left) displays panoramic imagery from the Spirit and Opportunity rover missions, Hipparcos Explorer (center) visualizes 120,000 stars within 500 light years from the earth, and Photon-Beam Tomography (right) is a volume rendering of the head of a cricket from a dataset of X-ray images.

6. Results and conclusions

<u>Results</u>

   Although the native display resolution is higher, the Personal Varrier effectively can display 640 x 800 autostereo over a 50° field of view at interactive frame rates. Cross-talk between the eyes (ghosting) is less than 5%, which is comparable to high-quality stereoscopic displays, except that Personal Varrier requires no glasses to be worn. Personal Varrier is orthographic: all dimensions are displayed in the same scale and no compression of the depth dimension occurs. Contrast is approximately 200:1. Personal Varrier is a small-scale VR display system with viewer-centered perspective, autostereo display, tether-less tracking, and real-time interactivity. A variety of distributed scientific visualization applications have been demonstrated using the CAVEwave network, including large data sets, viewpoint reconstruction, and real-time 3d teleconferencing.

<u>Limitations</u>

   Dynamic artifacts are visible during moderate-speed head movements, which disappear once motion has stopped. These are due to tracking errors and system latency. The current working range of the system is limited to 1.5 feet x 1.5 feet, which is restrictive. For example, a user who leans too far back in his or her chair will be out of range, producing erroneous imagery. However, the tracking system will re-acquire the user once he or she re-enters the usable range. Personal Varrier is currently a single-user system.

<u>Current and Future work</u>

   Camera-based tracking continues to evolve, striving to expand the working range and increase stability. A new user interface is being developed that will permit the user to self-calibrate the tracking system, without requiring a trained operator's assistance. New grid-based applications continue to be developed. Finally, new form factors are being investigated, for example, a cluster of 3 large displays is being considered as a larger desktop system. One screen may serve as a 2d repository for application toolbars, console, and control panel while the other two screens would be fitted with Varrier parallax barriers to become 3d displays. For example, one 3d display screen can be used for 3d data visualization while the other can serve as a collaborative 3d teleconferencing station.

7. Acknowledgement

8. References

[1] J. Ge, D. Sandin, T. Peterka, T. Margolis, T. DeFanti, Camera Based Automatic Calibration for the Varrier[TM] System, Proceedings of IEEE International Workshop on Projector-Camera Systems (PROCAM 2005), San Diego, 2005.

[2] J. Girado, Real-Time 3d Head Position Tracker System With Stereo Cameras Using A Face Recognition Neural Network, PhD thesis, University of Illinois at Chicago, 2004.

[3] J. Girado, D. Sandin, T. DeFanti, L. Wolf, Real-time Camera-based Face Detection using a Modified LAMSTAR Neural Network System, Proceedings of IS&T/SPIE's 15th Annual Symposium Electronic Imaging 2003, Applications of Artificial Neural Networks in Image Processing VIII, San Jose, California, 2003, pp. 20-24.

[4] D. He, F. Liu, D. Pape, G. Dawe, D. Sandin, Video-Based Measurement of System Latency, International Immersive Projection Technology Workshop 2000, Ames, Iowa, 2000.

[5] L. Lipton, M. Feldman, A New Autostereoscopic Display Technology: The SynthaGram, Proceedings of SPIE Photonics West 2002: Electronic Imaging, San Jose, California, 2002.

[6] T. Okoshi, Three Dimensional Imaging Techniques, Academic Press, N.Y., 1976.

[7] K. Perlin, S. Paxia, J. Kollin, An Autostereoscopic Display, Proceedings of ACM SIGGRAPH 2000, ACM Press / ACM SIGGRAPH, New York. Computer Graphics Proceedings, Annual Conference Series, ACM, 2000, 319-326.

[8] K. Perlin, C. Poultney, J. Kollin, D. Kristjansson, S. Paxia, Recent Advances in the NYU Autostereoscopic Display, Proceedings of SPIE, vol. 4297 (2001), San Jose, California.

[9] S. Rusinkiewicz, M. Levoy. Qsplat, A Multiresolution Point Rendering System for Large Meshes, Proceedings of ACM SIGGRAPH 2000, ACM Press / ACM SIGGRAPH, New York. Computer Graphics Proceedings, Annual Conference Series, ACM, 2000, p 343–352.

[10] D. Sandin, T. Margolis, G. Dawe, J. Leigh, T. DeFanti, The Varrier Autostereographic Display, Proceedings of SPIE, vol. 4297 (2001), San Jose, California.

[11] D. Sandin, T. Margolis, J. Ge, J. Girado, T. Peterka, T. DeFanti, The Varrier [TM] autostereoscopic virtual reality display, ACM Transactions on Graphics, Proceedings of ACM SIGGRAPH 2005 24, no. 3: (2005) 894-903.

[12] A. Schmidt, A. Grasnick, Multi-viewpoint Autostereoscopic Displays from 4D-Vision, Proceedings of SPIE Photonics West 2002: Electronic Imaging, San Jose, California, 2002.

[13] R. Singh, J. Leigh, T. DeFanti, F. Karayannis. TeraVision: a high resolution graphics streaming device for amplified collaboration environments, Future Generation Computer Systems vol. 19, no. 6, 2003, p 957-971.

[14] J.-Y. Son, S.A. Shestak, S.-S. Kim, Y.-J. Choi, Desktop Autostereoscopic Display with Head Tracking Capability, Proceedings of SPIE Vol. 4297 (2001), Stereoscopic Displays and Virtual Reality Systems VIII, San Jose, California.

[15] M. Thiebaux, H. Tangmunarunkit, K. Czajkowski, C. Kesselman, Scalable Grid-Based Visualization Framework. Technical report ISI-TR-2004-592, USC/Information Sciences Institute, June 2004.

[16] C. van Berkel, Image Preparation for 3D-LCD, Proceedings of SPIE Vol. 3639 Stereoscopic Displays and Virtual Reality Systems VI (1999), San Jose, California.

[17] Y. Wang, F. DeCarlo, D. Mancini, I. McNulty, B. Tieman, J. Bresnehan, I. Foster, J. Insley, P. Lane, G. von Laszewski, C. Kesselman, M. Su, and M. Thiebaux, A High-Throughput X-Ray Microtomography System at the Advanced Photon Source, Review of Scientific Instruments, 72(4), April 2001.

[18] D.F. Winnek, U.S. patent number 3,409,351, 1968.