Bunker View: Limited-range head-motion-parallax visualization for complex data sets

Andrei State, Suresh Balu and Henry Fuchs

Department of Computer Science
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina, 27514

ABSTRACT

This work presents a head-motion-parallax visualization system for arbitrarily complex computer graphics databases (such as biomedical data sets) using precomputed images. It features a large rear-projection screen and a blind restricting the range of possible vantage points. We describe implementation details for the high-resolution low-lag prototype system and relate first reactions from users.

1. INTRODUCTION

Visualizing complicated structures with spatially complex geometry, typically represented by very large numbers of graphics primitives (whether voxels, polygons, or others) is generally problematic and unsatisfactory on today's graphics engines. So far the spectrum of available visualization tools has been restricted to:

- precomputed cineloops (possibly enhanced by stereo displays) with oscillatory viewpoint motion (rocking)

- precomputed sequences of frames with arbitrary but pre-specified viewpoints (e. g., "flythroughs").

Today's graphics engines do not have enough rendering power to generate frames at rates that are high enough for the user to experience the 3D cue of head-motion-parallax. In particular, the frame rates are often too low for convincing head-tracked visualization. High image-generation lag further degrades the quality of such visualizations, especially in augmented reality scenarios where accurate registration between real and synthetic elements is essential.

2. PREVIOUS WORK

Previous visualization experiments attempting to combine stereoscopic viewing, head tracking, and high-resolution imagery (generated in real-time or precomputed) include the "CAVE,"[1] a virtual-reality system using several rear-projection screens driven by high-performance graphics engines; the screens form a cuboid; the user can move around within the cuboid. Head tracking is used to derive perspective projection parameters for the images presented on each screen; the real-time interlaced stereo imagery creates the perception of being surrounded by the 3D visualization, parts of which can appear to be inside the CAVE.

Deering described a CRT-based virtual reality system;[2] issues such as screen curvature, faceplate glass refraction and human eye parameters are discussed, together with their influence on virtual image generation and accurate virtual/physical world registration. Deering also suggests "hologram-like effects" could be achieved by using multiple precomputed views from densely sampled viewpoints. He concludes: "The ability of users to move their heads enables even naive users to rapidly perceive the spatial relationships of complex 3D shapes in a natural way."

Both the CAVE and Deering's system use real-time image generation systems which introduce significant lag between the user's head movement and the update of the display. Deering uses predictive tracking and small databases to reduce the effects of lag. For very large databases the image generation lag is unacceptably high, even in high-performance graphics engines, and makes virtual-reality type visualizations difficult or impossible to create, even with predictive tracking.

3. THE BUNKER VIEW SYSTEM

We propose to use image precomputation to avoid high lag and low frame rates for arbitrarily complex computer graphics models. Our system uses high-resolution stereoscopic visualization enhanced by limited-range head tracking. The essential technique is to restrict the user to a predefined range of viewpoints for which the system stores precomputed views. Whenever the position of the user's head changes, the system selects from the set of precomputed views a pair of stereoscopic images corresponding to the current vantage point and presents it to the user. Thus, in contrast to rocking cineloops or precomputed flythrough sequences, the display changes *in response* to the user's head motion.

3.1. Design considerations

The design of our prototype system was driven by the need to have storage capacity for a large number of precomputed images retrievable at high rates and with minimal lag (preferably comparable or inferior to the tracking system's own measurement delay; these delays are typically equal to or higher than 10 msec for a number of tracking systems examined in the literature[3]). If we assume a simple system without real-time image decompression, then the total number of precomputed images that can be used for a visualization session is determined by the available amount of memory in the system and the amount of memory required to store one frame.

The locations of the viewpoints used for image precomputation can be anything from an arbitrary curve in 3-space to an arbitrarily shaped 3D volume. We chose a 1D parametric space represented by a horizontal line segment located in front of the display screen. In contrast to unrestricted virtual reality systems such as the ones described above, our system must prevent the user from viewing the virtual image from vantage points that lie outside the precomputed range. To accomplish this, we have considered:

- blanking out the display whenever the user's head leaves the set of spatial locations of the precomputation vantage points (this was judged to be irritating since it would leave the user without a clue as to which way to move in order to make the imagery reappear), and

- physically constraining the user in such a way that the user's head cannot leave the set of locations of the precomputation vantage points (this was judged to be uncomfortable).

The eventually adopted solution consists of a vertical blind positioned parallel to the display screen (Figure 1); the blind blocks the user's view of the display (and thus the virtual image) except through a horizontal slit corresponding to the linear array of viewpoints. This effectively accomplishes the desired effect without unduly inconveniencing the user. While it is true that the user is free to move away from the slit, the vertical field of view diminishes rapidly in such cases; the arrangement seems to tend to "attract" the user toward the slit. This layout gave the system its name.

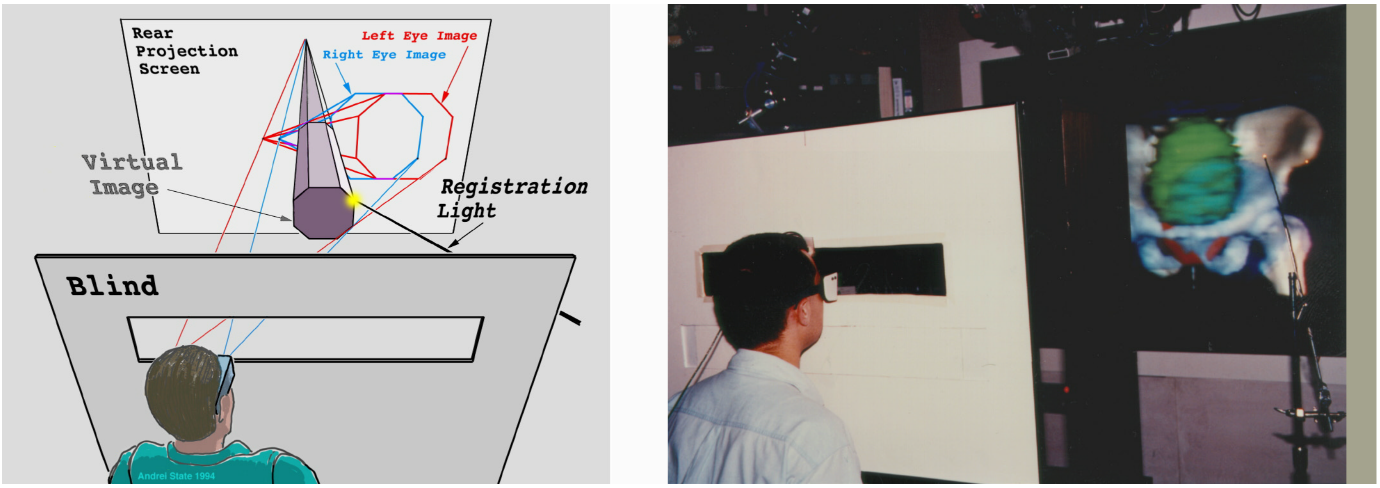
3.2. System components

Our prototype system has three major components:

- An image server; we are using the UNC-designed and -built Pixel-Planes 5 graphics engine[4] for this purpose because of the large amount of memory available in its computing nodes and the fast data paths connecting the (distributed) memory with the system's frame buffers. We decided on an image format adequate for field-sequential stereo display at 512-by-640 frame resolution or 256-by-640 resolution for a single field. 240 such single-eye 24-bit RGB images at 256-by-640 resolution (uncompressed) can be stored in the system. The display and update rates are both 72 Hz.

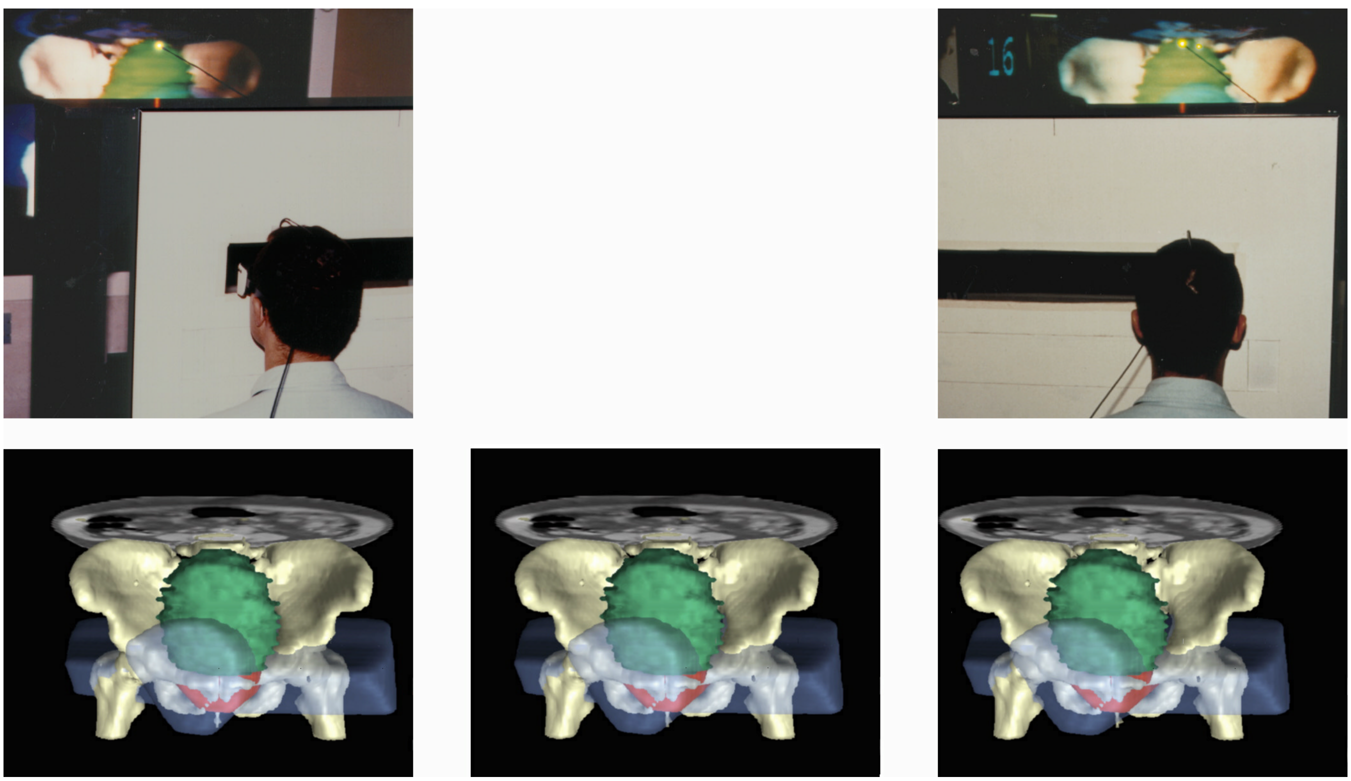
- A display subsystem consisting of a video projector equipped with a modified off-the-shelf stereo controller (StereoGraphics™) and a rear-projection screen. The field frequency is 144 Hz, twice the frame rate.

- A viewing/tracking subsystem consisting of a blind with a horizontal viewing slit and off-the-shelf active shutter LCD stereo polarizing glasses (StereoGraphics™) synchronized with the field rate via infrared signals issued by the display controller. The glasses are fitted with a sensor for an off-the-shelf electromagnetic tracking system (Polhemus™). The viewing slit is 700 mm wide; thus, with 240 possible viewpoints, the spacing between viewpoints is approximately 3 mm.



(a) (b)

Figure 1: (a) Sketch and (b) photo of Bunker View setup. Note mobile arm with miniature light bulb for registration verification.



(a) (b) (c)

Figure 2: (a) User viewing Bunker View imagery from the extreme left vantage point (top) and what the user sees (bottom). (b) Bunker View image seen from the center of the viewing slit. (c) User viewing Bunker View imagery from the extreme right vantage point (top) and what the user sees (bottom).

3.3. Image generation and display

The geometric layout of the observation area shown in Figure 1 (slit position and size, rear projection screen position and size, tracker source location, etc.) has been accurately measured. These parameters, together with the desired position, orientation, and size of the data set, are used to render a set of images by projecting the data set (located somewhere in the space beyond the viewing slit) onto the plane of the projection screen, from each of the 240 viewpoints located at 3 mm intervals along the viewing slit. Both feed-forward rendering algorithms (z-buffer, splatting) and feed-backward algorithms (ray casting, ray tracing) can be used, as long as the renderer supports oblique or off-center perspective projections. We modified existing (Pixel-Planes 5 based) volume rendering software (*vol2*[5]) for this purpose.

Figure 2a shows a user observing the display from the left endpoint of the viewing slit (top half). The bottom half of Figure 2a is one of the 2 stereo images the user sees from that vantage point. Figure 2c contains the equivalent views for the extreme right vantage point. Figure 2b is a view from the center of the viewing slit. The display images show an anatomical data set of a prostate cancer patient's pelvic region (a study resulting from the VISTAnet project[6]); the green structure is the urinary bladder; the translucent blue structure is a radiation dose isosurface resulting from simulated radiation therapy (one of the principal motivations for this study was to facilitate the comprehension of the 3D structure of such complex data sets containing multiple translucent, interpenetrating isosurfaces). A complete set of 240 images for all viewpoints along the slit can be computed in less than an hour on Pixel-Planes 5, longer on conventional workstations. Note that only a linear array of images is needed; each image can be used as a left eye view or a right eye view. The user's interocular distance is adjustable at viewing time, but must be a multiple of the distance between viewpoints (3 mm in our case). For example, any 2 of the three bottom images in Figure 2 can be used as a stereo pair, which some people can easily fuse into a 3D virtual image.

At viewing time the user sees a static 3D virtual image of the data set. At image precomputation time, this image can theoretically be positioned anywhere in space beyond the viewing slit, i. e., partly or completely in front or behind the display screen (for example, in Figure 1a the dataset is a very long octogonal prism that extends from halfway between blind and projection screen towards the horizon). However, according to previous research,[7] stereo parallax (i. e., *stereo separation*, not to be confused with head-motion-parallax) for comfortable stereo viewing should not exceed 5% of the display screen width or 85 mm for our 1700 mm wide display screen. Since human interocular distance rarely exceeds 85 mm, it follows that virtual image elements may be positioned anywhere from halfway between viewing slit and display screen to infinity without violating the 5% rule. So far we have only experimented with data sets positioned either half in front and half behind the projection screen, or mostly in front of the projection screen for a more powerful stereo effect, but without ever exceeding 5% negative stereo parallax.

Again following the advice in the referenced work,[7] we have also positioned our data sets in such a way as to avoid viewing situations in which virtual image elements with negative stereo parallax (i. e., appearing in front of the display screen) are clipped by any of the display screen edges (for all viewpoints along the slit). This avoids the visual confusion resulting from viewing situations in which obscuration and depth provide conflicting 3D cues.

3.4. Calibration and registration

Our prototype system exploits only a single (translational) component (parallel to the viewing slit) from the 6D tracker readings. We have implemented a translation table to calibrate for position reporting nonlinearities (field distortion), but were unable to find any significant ones in the system we were using and within its restricted range of operation (i. e., along the 700 mm wide viewing slit). The field distortions inherent in some electromagnetic trackers are non-negligible for large tracking areas. On the other hand, the smaller the visualization setup (including display screen) is, the more precise the required tracking and display calibration required. Our system avoids many of the problems described by Deering[2] simply by using a large, flat display screen located much farther away from the user than a desktop monitor, and by showing a large virtual image whose size is comparable to the size of the display screen. Thus, in contrast to Deering's desktop system, small inaccuracies in tracker readings do not result in noticeable display distortion, and we have so far neglected issues such as display screen curvature (our rear-projection screen is completely flat) or the precise position of the center of projection within the eye (the observer's eye is small compared to the size of the virtual image).

To verify how accurately virtual/physical world registration in our system is, we used a miniature light bulb mounted on a mobile arm (visible in Figure 1). The bulb was positioned in the space between viewing slit and screen such that it appeared to coincide with a particular feature (of negative stereo parallax) in the image when viewed from the extreme left viewpoint (for example, in Figure 1a, the registration light is positioned on one of the vertices of the octogonal prism). Then it was verified that while the observer moved towards the extreme right vantage point and subsequently back and forth along the slit, the bulb would continue to appear attached to the same feature in virtual image space. Figures 2a and 2c show the bulb positioned at the top of the green urinary bladder structure—if one assumes the camera's point of view coincides with the observer's, which is obviously not true in our demonstration images (a better way to demonstrate this, aside from operating the actual system, is with a video recording made from a tracked camera as in Deering's SIGGRAPH '92 presentation).

4. VIEWING BUNKER VIEW IMAGERY

While noting the lack of "swimming," the quick response and the accurate registration to real-world objects, several observers have also reported a slight "jumping" effect when viewing Bunker View imagery. Our current hypothesis is that this is related to the "it seems to move more than it should according to my head motion" complaint we have also heard occasionally. We suspect it is due to the currently still very coarse spacing between viewpoints (3 mm): at any moment in time while the observer moves along the slit, the displayed image either doesn't change at all (if the observer hasn't moved far enough yet), or it jumps quickly as if the observer had instantaneously moved by 3 mm. We currently speculate that some observers may notice the jumping as such, while others may subjectively integrate it into what they report as apparently exaggerated response.

5. CONCLUSIONS

Bunker View, a variation of "fish tank virtual reality," is situated somewhere in between "simple" precomputed imagery (such as flythrough visualizations or rocking cineloops) and unrestricted head-motion-parallax (such as in virtual reality systems). It combines some of the benefits (achieved or desired) and some of the detriments of both techniques:

- High image quality, uncompromised by real-time computational needs
- Head-motion-parallax, albeit spatially restricted
- Limited, predefined range of viewpoints
- High, constant display update rates
- Low, constant lag (independent of database size), bounded by tracking system lag and display update rate.
- Excellent registration between synthetic imagery and real world objects

Bunker-View-type systems are not only useful as mechanisms for user studies, but also as pedagogical tools that can clearly demonstrate (due to their potential to provide a "gold-standard" for see-through virtual reality) the visualization quality that we hope to achieve in the future with systems containing real-time image generators.

6. FUTURE WORK

We plan to improve the quality of the Bunker View visualization by making it "rock-solid"; based on our current understanding, this means:

- Increasing the number of precomputed views and reducing viewpoint spacing
- Reducing the system's end-to-end lag even further, possibly by using optical tracking and/or predictive methods (it has been demonstrated that systems with low, constant lag are excellent candidates for predictive methods[8]).

We also plan to supplement our anecdotal evidence by systematic user studies that will attempt to answer questions such as:

- Can the "jumping" and "excessive motion" complaints be eliminated by very dense viewpoint spacing?
- How much better is Bunker View than precomputed imagery techniques for a specific visualization task?
- How much worse than unrestricted head-motion-parallax is it?

Other possible areas of investigation are:

- Experiment with different viewing range and display configurations
- Take head rotations into account (so far only translations are exploited)
- Add interaction to the system: for example, introduce moving/changing elements into the visualization by combining precomputed imagery with low-lag real-time image synthesis

As computer system memory sizes increase and fast image decompression techniques become available in multimedia-oriented machines, it will become possible to realize Bunker-View-type visualizations on desktop machines. Such systems may cache a small number of images in decompressed form in local memory, and use look-ahead techniques to request from fast mass storage units the images expected to be displayed in the immediate future. Finally, as high-bandwidth communications networks are deployed, remote image servers could be accessed in this way.

7. ACKNOWLEDGMENTS

We thank Gary Bishop and Stephen M. Pizer for their significant contributions and constructive criticism, as well as Jannick Rolland for her critical review of this paper. We also thank the anonymous reviewers for their helpful comments and constructive criticism.

This work was supported by NSF and ARPA under Cooperative Agreement NCR-8919038 with CNRI ("VISTAnet: A Very High Bandwidth Prototype Network for Interactive 3D Imaging"), by BellSouth, by GTE and by ARPA ISTO contract DAEA 18-90-C-0044 ("Advanced Technology for Portable Personal Visualization").

8. REFERENCES

[1] Cruz-Neira, Carolina, Daniel J. Sandin, Thomas A. DeFanti, "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE," *Computer Graphics*, (Proceedings of SIGGRAPH '93), pp. 135-142, July 1993.

[2] Deering, Michael, "High-Resolution Virtual Reality," *Computer Graphics*, 26(2) (Proceedings of SIGGRAPH '92), pp. 195-202, July 1992.

[3] Mine, Mark R., "Characterization of End-to-End Delays in Head-Mounted Display Systems," *UNC-CS Technical Report* TR93-001, March 1993.

[4] Fuchs, Henry et al., "Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor-Enhanced Memories," *Computer Graphics*, 23(3) (Proceedings of SIGGRAPH '89), pp. 79-88, July 1989.

[5] Neumann, Ulrich, Andrei State, Hong Chen, Henry Fuchs, Tim J. Cullip, Qin Fang, Matt Lavoie, John Rhoades, "Interactive Multimodal Volume Visualization for a Distributed Radiation-Treatment Planning Simulator, " *UNC-CS Technical Report TR94-040,* June 1994.

[6] State, Andrei, Julian Rosenman, MD, Henry Fuchs, Tim J. Cullip and Jim Symon, "VISTAnet: Radiation therapy treatment planning through rapid dose calculation and interactive 3D volume visualization," these proceedings.

[7] Akka, Robert, "Automatic software control of display parameters for stereoscopic graphics images," *SPIE*, Vol. 1669 (1992), pp. 31-38, 1992.

[8] Azuma, Ron, "Improving Static and Dynamic Registration in an Optical See-through HMD," *Computer Graphics* (Proceedings of SIGGRAPH '94), July 1994.