0.0: Real-Time Natural 3D Content Displaying with HoloVizio displays

Péter Tamás Kovács, Zoltán Gaál, Attila Barsi, Zoltán Megyesi
Holografika Ltd., Budapest, Hungary

Abstract
In this paper we present HoloVizio technology as a solution to real time natural content displaying. The presented display technology is capable of reproducing large field of view continuous parallax light fields, providing realistic 3D experience of displayed scenes and objects. In this paper we demonstrate that the system is capable of real time displaying of dense light field streams originating from multiple cameras. We present the results of a real time acquisition - displaying solution using the HoloVizio system and a camera array of 27 cameras.
In this paper we summarize the technology, describe the components of the system and show results and measurements on natural scenes.

1. Introduction
Displaying 3D images is a major step towards realistic visualization. Providing perceivable 3D information can be crucial in applications of many areas including architecture design, biological research, industrial engineering, geological surveying, medical visualization and more. Several technologies addressing this task are available. These include stereoscopic displays with or without tracking, multiview systems, volumetric and holographic systems.

The mainstream solutions in 3D displaying are active or passive stereo glasses, which have the advantages of being cheap and easily available. However, they as most stereoscopic systems can only provide 3D view for a single fixed position. Autostereoscopic displays are capable of providing 3D images to multiple positions, but the most widespread displays (lenticular or parallax barrier systems) have limited 3D ray count (3D resolution), provide continuous view for a narrow FOV, and viewers moving may experience jumps when leaving or entering valid zones. The HoloVizio technology is capable of providing continuous parallax 3D images simultaneously for a wide viewing zone.

The HoloVizio displays are capable of displaying high quality horizontal parallax light fields. However, a dense, large FOV light field contains much information and consequently the handling of such amount of data in real-time is challenging. The displaying of large FOV light field videos is a challenging task, which is implemented using a small PC cluster. Also this data is hard to acquire, store or transfer in real time. We connected a 10 Million light-ray HoloVizio system (HoloVizio 240P) with a 27 camera array capturing 18 Million light rays. The captured natural content was transferred, processed and displayed in real time. The same system was later used to display live content on our large-scale display (HoloVizio 720RC).

2. HoloVizio Technology
The patented HoloVizio technology uses a different approach from stereoscopic, multiview, volumetric and holographic systems. It uses a specially arranged array of optical modules and a holographic screen. Each point of the holographic screen emits light beams of different color and intensity to the various directions. The light beams generated in the optical modules hit the screen points in various angles and the holographic screen makes the necessary optical transformation to compose these beams into a perfectly continuous 3D view. With proper software control, light beams leaving the pixels propagate in multiple directions, as if they were emitted from the points of 3D objects at fixed spatial locations [1][2] (see).

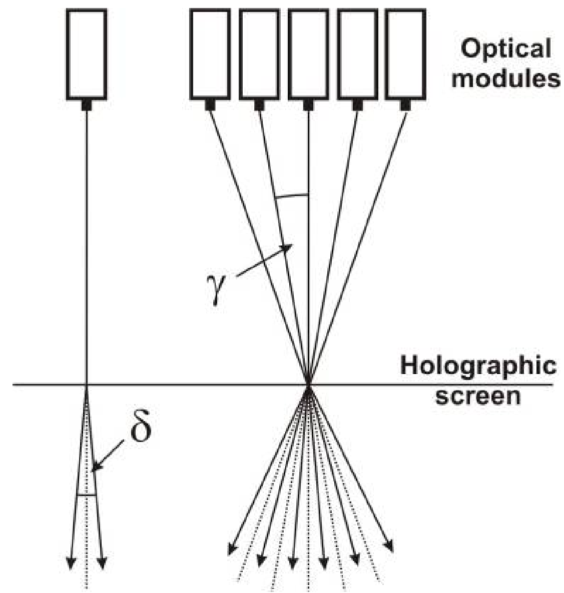

Figure 1: HoloVizio principle

2.1. HoloVizio Displays
The HoloVizio displays are available in different sizes ranging from desktop displays to large scale systems. Beside the screen size, the most important factor is the number of light rays emitted from the display, which can be considered the 3D resolution of the display. (The rays can be considered as the “pixels” of the system).

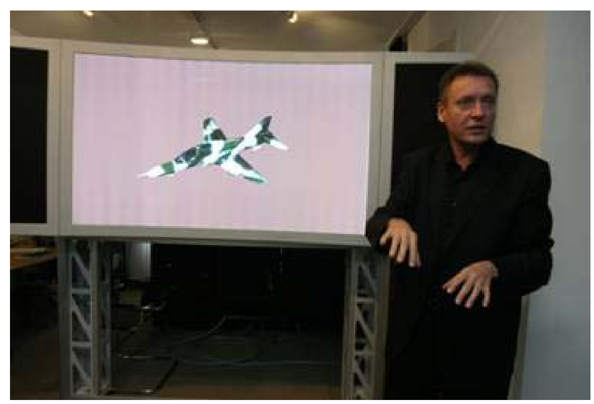

Figure 2: 50 MPixel HoloVizio display

A 50 MPixel large-scale system [3] (see) has been developed with a screen diagonal above 1.8m. The optical system consists of compact projection modules, arranged in horizontal rows. The system has a high angular resolution; approximately 50 independent light beams originate from each pixel. A PC-based render cluster feeds the display with 50 MPixel in real-time and a sophisticated control system controls the projectors, PCs, the network, power supplies and monitors all system parameters.

With the HoloVizio approach it is possible to build displays that have excellent image resolution of 1920x1080 or beyond, large FOV above 100 degrees, large Field-of-Depth, and at the same time the number of pixels being in the range of hundreds of millions.

As for monitor style 3D displays a 32", 10 MPixel 16:9 aspect model [4] is currently available (Fig 3). This model is in the dimensions of normal TV sets. The image is provided by 128 microdisplays, and controlled by 1-4 PCs.

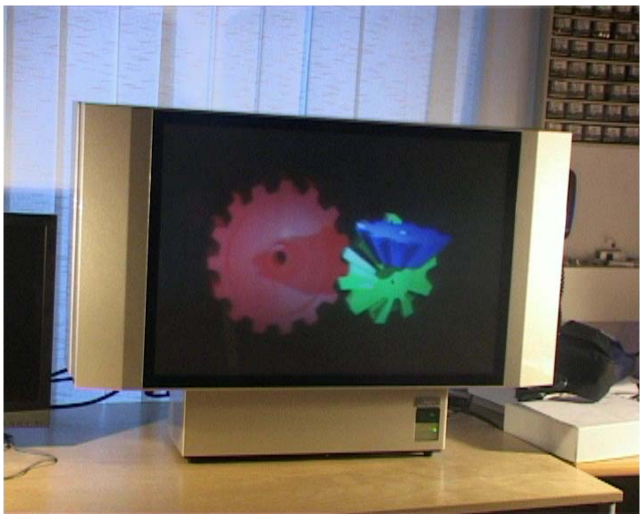

Figure 3: 10 MPixel HoloVizio display

The model in between is the HoloVizio 240P display (Fig 4), the first HoloVizio featuring a slim optical design, which, despite being a projection based display, allows it to be only 70 cm deep, and is controlled by just 3 built-in PCs.

Figure 4: HoloVizio 240P winning the Second Prize of the Best Exhibit Award at ICT Lyon [14]

2.2. Image Generation for HoloVizio Displays

There are several possibilities for displaying 3D data on the HoloVizio: interfacing interactive graphics applications to the holographic displays, displaying light-field data, and creating images directly for the HoloVizio lighting modules.

For interactive graphics application, a wrapper library intercepts all calls and related the application sends to OpenGL and transfers them through the network for replaying on the rendering nodes inside the HoloVizio. The OpenGL wrapper supports most of the usual features of the OpenGL 2.0 standard, including user-defined GLSL, ARB, Cg and MetaSL shaders, and VBOs. It has been successfully used with various industrial applications (see Figures 5 & 6), and is compatible with applications based on common OpenGL-based visualization libraries such as: OpenInventor, Coin3D, OpenSceneGraph, AVS/Express.

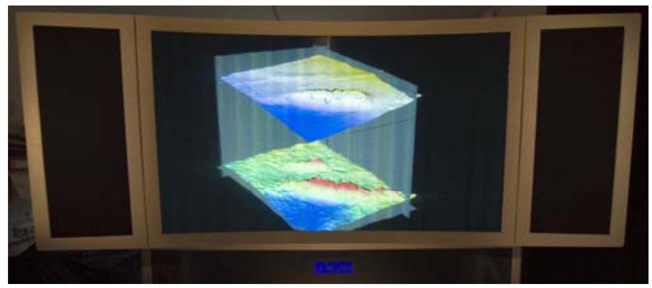

Figure 5. HoloVizio running an Oil&Gas application

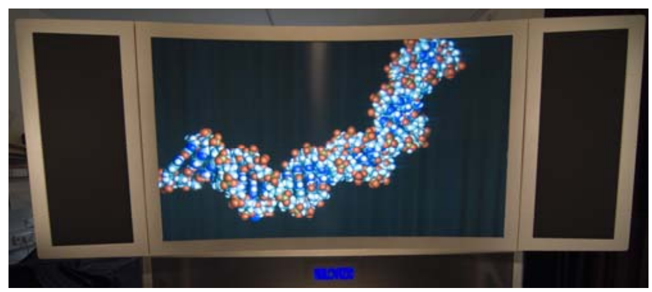

Figure 6. HoloVizio running Visual Molecular Dynamics

For displaying other content (eg. pre-rendered animations), a 3D converter and video player application is provided including the necessary plug-ins for popular 3D-renderer software. This rendering-conversion pipeline is capable of creating high quality content for the HoloVizio (see Figure 7).

Figure 7. HoloVizio showing a rendered car model

3. Real-time displaying of dense light field

The existing software system concentrated on displaying synthetic content, which was the dominant use case of existing HoloVizio displays. Although this is sufficient for professional applications that are working with 3D data, displaying live 3D content is a major step in the direction of widespread use such as 3D Cinema or 3DTV.

3.1. Light-field Content

Unfortunately, light-field content is very rare (with some exceptions [6][7]). Using multi-view content – which is also quite scarce – provides suboptimal viewing experience on HoloVizios both in terms of very narrow FOV and angular resolution compared to the capabilities of the displays, even if the latter can be somewhat compensated by using Depth Based Rendering, Image Based Rendering, or a hybrid approach [8].

Although several companies offer Time Freeze shooting [9], these are targeting the visual effect and movie industry, where real-time acquisition, transmission and playback is not a target and is also costly.

3.2. State Of The Art

Several results describing real-time light-field capture and display have been published in recent years, achieving significant advances. The random access light-field camera provides very good results with 2D or stereo displays due to selective transmission [10], but as the authors pointed out, this approach is less applicable with 3D displays which typically need access to all light-field data. The TransCAIP system [11] uses a single PC and GPU algorithms to achieve interactive speeds with an impressive number of cameras, however, our system provides better results in terms of resolution, frame rate, and better angular resolution, moreover, our system is designed to be highly scalable. The impressive MERL 3D TV System [12] uses a symmetrical system with a high number of PCs and a lenticular-lens based display to create live 3D visuals, however the HoloVizios we used have far better 3D image quality compared to their 3D display, moreover they use excessive number of PCs for capturing, processing and rendering.

3.3. 3D acquisition system and calibration

We built a camera system with 27 USB CCD cameras and connected them to 3 PC computers. The cameras were in an evenly spaced linear arrangement, and each one captured with 640x480@15 FPS or 960x720@10 FPS resolutions (see Fig 8).

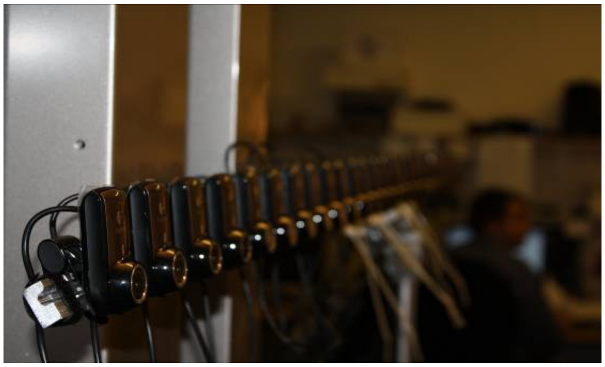

Figure 8: 3D Camera System

The capture system was directly connected through a single gigabit Ethernet connection to the demonstration 3D display, supported by a 3 PC computation cluster, which are an integral part of the HoloVizio display. The camera system was calibrated off line using a semi-automatic calibration method, using images of a previously known reference object [15,16]. As a result of the calibration, the camera light field parameters were estimated based on intrinsic and extrinsic camera calibration parameters, (See Figure 9). These estimates are further refined to minimize the error of the estimated model, resulting in very good 3D image quality.

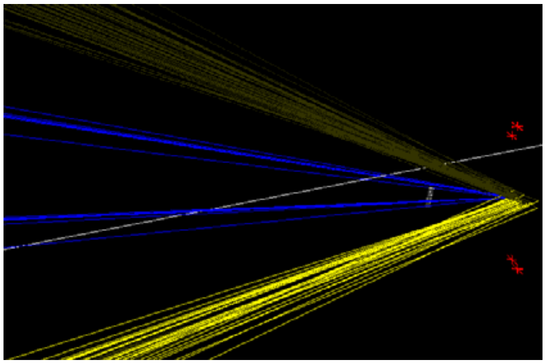

Figure 9: Estimated Light Field before refinement

3.4. Light-field conversion

For the task of dense light field displaying a massively parallel reordering and filtering of the original camera images is required. We were utilizing both CPU and GPU threads for this task. The original MJPEG images are arriving to the cluster's nodes on a single Gigabit Ethernet with approx 10% link utilization. Each individual channel has its own CPU thread that decodes the Huffman encoding and does the inverse DCT algorithm for the incoming JPEG image. This yields a YUV image on the CPU. (IDCT has also been implemented on the GPU, but we found that it was not the real bottleneck.) On the GPU we do the light-field conversion and reordering, filtering and the YUV-RGB conversion. We use OpenGL 3.0 shaders and 2D texture arrays to have an easy access to individual camera images. This yields the correct bilinear filtering. There is also a network-based synchronization scheme for displaying the final rendered images.

The current implementation's two performance-critical points are the Huffman decoding of the JPEG images and the upload speed on the PCI-Express bus, even though the cameras themselves cannot provide higher resolution with good frame rate.

4. Results and conclusion

To perform the time measurements, we used a 10Mpixel HoloVizio test system, but the 50 Mpixel display was also used for demonstration. This demonstration provides a practical possible use of the system in true 3D teleconferencing (see Figure 10). The reasons for using 3D in telepresence is twofold. 3D visualization provides a more life-like experience, and 3D light-field capturing and visualization inherently solves the problem of incorrect eye gaze, which is a serious drawback of all 2D teleconferencing systems, even high-end ones. Although the eye-gaze problem can be solved with multi-view displays [13], such a solution is limited to a fixed number of people, carefully positioned to the calibrated locations. Our solution allows people to move freely, experiencing a continuous motion parallax.

Figure 10: Continuous motion-parallax, true 3D teleconferencing

The 27 cameras captured 960x720 resolution video streams. The image data was converted on-the-fly to light field format and the continuous parallax 3D image stream was displayed.

Two PC clusters of 3 PCs (built from commercial components) were used, one on the acquisition and one on the display side. With this setup, we reached 15 frames per seconds playback speed for the 640x480 resolution stream and 10 frames per second for the 960x720 resolution stream. The bottleneck here was the camera acquisition speed.

5. Future Work

Future improvements are planned for building a system with a higher number of cameras with higher resolution and frame rate. A self-contained intelligent camera hardware eliminating PCs is under development, providing 64 2MPixel streams at 30 FPS, synchronized.

As the image format used at the moment is suboptimal, not being compact enough for Internet transmission, a two-layered approach for decoding and rendering is being developed for using different encoding during transmission (more compact), and light-field rendering (more GPU-friendly), which also allows arbitrary number of cameras and arbitrary resolution and frame rate usage . Full GPU decoding of traditional video codecs is also under investigation.

6. Acknowledgements

The development of the 10Mpixel test display system and camera system has been supported by EU IST-FP6 Integrated Project OSIRIS (IST-33799 IP) [5]. The FP6 project OSIRIS [5] aims to create novel display systems including a high resolution LED based compact display that is capable of real-time playback of live captured natural content. In the same project HoloVizio technology is applied to create a 3D Cinema application.

7. References

[1] T. Balogh: Method and apparatus for displaying three-dimensional images, U.S. Patent 6,201,565, EP 0900501, Feb 04, 1997

[2] T. Balogh: Method and apparatus for displaying 3d images, U.S. Patent Appl. 20030156077, EP 1 285304, May 19, 2000

[3] T. Balogh, T. Forgacs, T. Agocs, E. Bouvier, F. Bettio, E. Gobbetti and G. Zanetti: A Large Scale Interactive Holographic Display, IEEE VR2006 Conference, Virginia, USA

[4] T. Balogh et al: A Scalable Hardware and Software System for the Holographic Display of Interactive Graphic Applications, EuroGraphics 2005, Dublin

[5] OSIRIS Project – Original System for Image Rendition via Innovative Screens, EU IST-FP6 Integrated Project OSIRIS (IST-33799 IP), http://www.osiris-project.eu/

[6] UCSD/MERL Light Field Repository, http://vision.ucsd.edu/datasets/lfarchive/

[7] The (New) Stanford Light Field Archive, http://lightfield.stanford.edu/

[8] Z.Megyesi, A. Barsi, T. Balogh: 3D Video Visualization On The Holovizio™ System, 3DTV-CON 2008, Istanbul, Turkey

[9] Upstage Productions: Time Freeze, http://www.upstage.com.au/content/standard.asp?name=Time_freeze

[10] J. C. Yang, M. Everett, C. Buehler, L. McMillan: A Real-Time Distributed Light Field Camera, 13th Eurographics Workshop on Rendering, 2002

[11] Y. Taguchi, T. Koike, K. Takahashi, T. Naemura: TransCAIP: Live Transmission of Light Field from a Camera Array to an Integral Photography Display, ACM SIGGRAPH ASIA 2008

[12] W. Matusik, H. Pfister: 3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes, ACM SIGGRAPH 2004

[13] P. Lincoln, A. Nashel, A. Ilie, H. Towles, G. Welch, H. Fuchs: Multi-view lenticular display for group teleconferencing, IMMERSCOM 2009

[14] European Commission, ICT Event 2008, http://ec.europa.eu/information_society/events/ict/2008

[15] Z. Zhang: A Flexible New Technique for Camera Calibration, Technical Report MSR-TR-98-71, Microsoft Research 1998

[16] R. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," IEEE J. Robotics and Automation, vol. 3, no. 4, pp. 323-344, Aug. 1987.