Interactive three-dimensional holographic displays: seeing the future in depth

Mark Lucente
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218, Yorktown Heights, New York 10598 USA
lucente@watson.ibm.com
www.research.ibm.com/people/l/lucente
(tel) 914-945-2980
(fax) 914-945-2141

1.0   Introduction

Computer graphics is confined chiefly to flat images. Images may look three-dimensional (3-D), and sometimes create the illusion of 3-D when displayed, for example, on a stereoscopic display [1-3]. Nevertheless, when viewing an image on most display systems, the human visual system (HVS) sees a flat plane of pixels. Volumetric displays can create a 3-D computer graphics image, but fail to provide many visual depth cues (e.g., shading, texture gradients) and cannot provide the powerful depth cue of overlap (occlusion). Discrete parallax displays (such as lenticular displays) promise to create 3-D images with all of the depth cues, but are limited by achievable resolution. Only a real-time electronic holographic ("holovideo") display [4-12] can create a truly 3-D computer graphics image with all of the depth cues (motion parallax, ocular accommodation, occlusion, etc.) and resolution sufficient to provide extreme realism [2]. Holovideo displays promise to enhance numerous applications in the creation and manipulation of information, including telepresence, education, medical imaging, interactive design, and scientific visualization.

The technology of electronic interactive three-dimensional holographic displays is in its first decade. Though fancied in popular science fiction, only recently have researchers created the first real holovideo systems by confronting the two basic requirements of electronic holography: (1) computational speed, and (2) high-bandwidth modulation of visible light. This article describes the approaches used to address these

problems, as well as emerging technologies and techniques that provide firm footing for the development of practical holovideo.

2.0 Electroholography Basics

Optical holography, used to create 3-D images, begins by using coherent light to record an interference pattern [13]. Illumination light is modulated by the recorded holographic fringe pattern (called a "fringe"), subsequently diffracting to form a 3-D image. As illustrated in Figure 1, a fringe region that contains a low spatial frequency component diffracts light by a small angle. A region that contains a high spatial frequency component diffracts light by a large angle. In general, a region of a fringe contains a variety of spatial frequency components and therefore diffracts light in a variety of directions.

An electroholographic display generates a 3-D holographic image from a 3- D description of a scene. This process involves many steps, grouped into two main processes: (1) computational, in which the 3-D description is converted into a holographic fringe, and (2) optical, in which light is modulated by the fringe. Figure 2 shows a map of the many techniques used in these two processes.

The difficulties in both fringe computation and optical modulation result from the enormous amount of information (or "bandwidth") required by holography. Instead of treating an image as a pixel array with a sample spacing of approximately 100 microns as is common in a two-dimensional (2-D) display, a holographic display must compute a holographic fringe with a s ample spacing of approximately 0.5 micron to cause modulated light to diffract and form a 3-D image.

A typical palm-sized full-parallax (light diffracts vertically as well as horizontally) hologram has a sample count (i.e., "space-bandwidth product" or simply "bandwidth") of over 100 gigasamples. Horizontal-parallax-only (HPO) imaging eliminates vertical parallax resulting in a bandwidth savings of over 100 times without greatly compromising display performance [8]. Holovideo is more difficult than 2-D displays by a factor of about 40,000, or about 400 for an HPO system. The first holovideo display created small (50 ml) images that required minutes of computation for each update [9]. New approaches, such as holographic bandwidth compression and faster digital hardware, enable computation at interactive rates and promise to continue to increase the speed and complexity of displayed holovideo images [5]. At

present, the largest holovideo system creates an image that is as large as a human hand (about one liter) [11]. Figure 3 shows typical images displayed on the MIT holovideo system.

3.0 Holographic Fringe Computation

The computational process in electroholography converts a 3-D description of an object or scene into a fringe pattern. Holovideo computation comprises two stages: (1) a computer graphics rendering-like stage, and (2) a holographic fringe generation stage in which 3-D image information is encoded in terms of the physics of optical diffraction. (See Figure 2.)

The computer graphics stage often involves spatially transforming polygons (or other primitives), lighting, occlusion processing, shading, and (in some cases) rendering to 2-D images. In some applications, this stage may be trivial. For example, MRI data may already exist as 3-D voxels, each with a color or other characteristic.

The fringe generation stage uses the results of the computer graphics stage to compute a huge 2-D holographic fringe. This stage is generally more computationally intensive, and often dictates the functions performed in the computer graphics stage. Furthermore, linking these two computing stages has prompted a variety of techniques. Holovideo computation can be classed into two basic approaches: interference-based and diffraction-specific.

3.1 The Interference-Based Approach

The conventional approach to computing fringes is to simulate optical interference, the physical process used to record optical holograms [13]. Typically, the computer graphics stage is a 3-D filling operation which generates a list of 3-D points (or other primitives), including information about color, lighting, shading, and occlusion.

Following basic laws of optical propagation, complex wavefronts from object elements are summed with a reference wavefront to calculate the interference fringe [8]. This summation is required at the many millions of fringe samples and for each image point, resulting in billions of computational steps for small simple holographic images. Furthermore, these are complex arithmetic operations involving trigonometric functions and square roots, necessitating expensive floating point calculations. Researchers using the

interference approach generally employ supercomputers and use simple images to achieve interactive display [8]. This approach produces an image with resolution that is finer than can be utilized by the human visual system.

Stereograms: A stereogram is a type of hologram that is composed of a series of discrete 2-D perspective views of the object scene [4]. An HPO stereogram produces a view-dependent image that presents in each horizontally displaced direction the corresponding perspective view of the object scene, much like a lenticular display or a parallax barrier display [1-3]. The computer graphics stage first generates a sequence of view images by moving the camera laterally in steps. These images are combined to generate a fringe for display.

The stereogram approach allows for computation at nearly interactive rates when implemented on specialized hardware [4]. One disadvantage of the stereogram approach is the need for a large number of perspective views to create a high-quality image free from sampling artifacts, limiting the computation speed. New techniques may improve image quality and computational ease of stereograms [14].

3.2 The Diffraction-Specific Approach

The diffraction-specific approach breaks from the traditional simulation of optical holographic interference by working backwards from the 3-D image [5-7]. The fringe is treated as being subsampled spatially (into functional holographic elements or "hogels") and spectrally (into an array of "hogel vectors"). One way to generate a hogel-vector array begins by rendering a series of orthographic projections, each corresponding to a spectral sample of the hogels. The orthographic projections provide a discrete sampling of space (pixels) and spectrum (projection direction). They are easily converted into a hogel-vector array [5]. A usable fringe is recovered from the hogel-vector representation during a decoding step employing a set of precomputed "basis fringes."

 The multiple-projection technique employs standard 3-D computer graphics rendering (similar to the stereogram approach). The diffraction-specific approach increases overall computation speed and achieves bandwidth compression. A reduction in bandwidth is accompanied by a loss in image sharpness -- an added blur that can be matched to the acuity of the HVS simply by choosing an appropriate compression ratio and sampling parameters. For a compression ratio (CR, the ratio between the size of the fringe and the

hogel-vector array) of 8:1 or lower, the added blur is invisible to the HVS. For CR of 16:1 or 32:1, good images are still achieved, with acceptable image degradation [5].

Specialized Hardware: Diffraction-specific fringe computation is fast enough for interactive holographic displays. Decoding is the slower step, requiring many multiplication-accumulation calculations (MACs). Specialized hardware can be utilized for these simple and regular calculations, resulting in tremendous speed improvements. Researchers using a small digital signal processing (DSP) card achieved a computation time of about one second for a 6-MB fringe with CR=32:1 [15]. In another demonstration, the decoding MACs are performed on the same Silicon Graphics RealityEngine2 (RE2) used to render the series of orthographic projections [5]. The orthographic projections rendered on the RE2 are converted into a hogel-vector array using filtering. The array is then decoded on the RE2, as shown in Figure 4. The texture-mapping function rapidly multiplies a component from each hogel vector by a replicated array of a single basis fringe. This operation is repeated several times, once for each hogel-vector component, accumulating the result in the accumulation buffer. A computation time of 0.9 seconds was achieved for fringes of 6-MB with CR=32:1 [5].

Fringelets: Fringelet bandwidth compression (Figure 2) further subsamples in the spatial domain [6]. Each hogel is encoded as a spatially smaller "fringelet." Using a simple sample-replication decoding scheme, fringelets provide the fastest method (to date) of fringe computation. Complex images have been generated in under one second for 6-MB fringes [6]. Furthermore, a "fringelet display" can optically decode fringelets to produce a CR-times greater image volume without increased electronic bandwidth.

4.0 Optical Modulation and Processing

The second process of a holographic display is optical modulation and processing. Information about the desired 3-D scene passes from electronic bits to photons by modulating light with a computed holographic fringe using spatial light modulators (SLMs). The challenge in a holographic display arises from the many millions of samples in a fringe. Successful approaches to holographic optical modulation exploit parallelism and/or the time-response of the HVS.

4.1 Liquid-Crystal and Related SLMs

A liquid crystal display (LCD) is a common electro-optic SLM used to modulate light for projection of 2-D images. A typical LCD contains about one million elements ("pixels"). A one-million-sample fringe can produce only a small flat image. A magneto-optic SLM, which uses the magneto-optic effect to electronically modulate light, often contains less than one million elements [16]. Early researchers using LCD SLMs or magneto-optic SLMs created small planar images [16-19]. The low pixel count of typical LCDs is overcome by tiling together several such modulators [12].

For any modulation technique, several issues must be addressed. Modulation elements are too big - typically 50 microns wide (in an LCD) compared to the fringe sampling pitch of about 0.5 micron. Demagnification is employed to reduce the effective sample size, with the necessary but unattractive effect of proportionally reducing the lateral dimensions of the image.(See Figure 5.) Holographic imaging may employ either amplitude or phase modulation. LCDs are basically phase modulators when used without polarizing optics. Phase modulation can be more optically efficient, and so is most often used. Finally, it is desirable to employ modulators possessing many levels of modulation, i.e., grayscale. Common LCDs have nominally 256 grayscale levels, sufficient for producing reasonably complex images.

Deformable micro-mirror devices (DMDs) are micromechanical SLMs fabricated on a semiconductor chip as an electronically addressed array of tiny mirror elements. Electrostatically depressing or tilting each element modulates the phase or amplitude of a reflected beam of light. A phase-modulating device was used to create a small flat holographic image [20], and a binary amplitude-modulating DMD was used to create a small interactive 3-D holographic image [21].

4.2 Scanned Acousto-Optic Modulator (AOM)

The time-multiplexing of a very fast AOM SLM has been used in holovideo. A wide-aperture AOM phase modulates only about 1000 samples at any one instant in time, using a rapidly propagating acoustic wave within a crystal2E By scanning the image of modulated light with a rapidly moving mirror, a much larger apparent fringe can be modulated. The latency of the HVS is typically 20 ms, and the eye time-integrates to see the entire fringe displayed during this time interval. This technique was invented and exploited by researchers at the MIT Media Laboratory to produce the world's first real-time 3-D holographic display in 1989 [9,10]. A generalized schematic of this approach is shown in Figure 6. After RF processing,

computed fringes traverse the aperture of an AOM (as acoustic waves), which phase-modulate a beam of laser light. Two lenses image and demagnify the diffracted light at a plane in front of the viewer. The horizontal scanning system angularly multiplexes the image of the modulated light. A vertical scanning mirror reflects diffracted light to the correct vertical position in the hologram plane.

One advantage of the scanned-AOM system is that it can be scaled up to produce larger images. The first images produced in this way were 50 ml, generated from 2-MB fringes [9]. More recently, by building a scanned-AOM system with 18 parallel modulation channels, images created from a 36-MB fringe occupy a volume greater than one liter [11].

One disadvantage of the scanned-AOM approach is the need to convert digitally computed fringes into high-frequency analog signals. The 18-channel synchronized high-speed framebuffer system used at MIT was made for this application, and was a major practical obstacle in this approach [15]. LCDs, DMDs, and other SLMs are more readily interfaced to digital electronics. Indeed, LCD SLMs are commonly constructed to plug directly into a digital computer, or are built on an integrated circuit chip [22]. Another disadvantage of the scanned-AOM approach is the need for optical processing. Typical LCD-based holographic displays require only demagnification and the optical concatenation of multiple devices. The time-multiplexing of the scanned-AOM system requires state-of-the-art scanning mirrors which must be synchronized to the fringe data stream. Despite these obstacles, the scanned-AOM approach has produced the largest holovideo images [6].

4.3 Other Techniques

Color: Full-color holovideo images are produced by computing three separate fringes. Each represents one of the additive primary colors (red, green, and blue) taking into account the three different wavelengths used in a color holovideo display. The three fringes are used to modulate three separate beams of light (one for each primary color) [10].

SAW AOM: Recently, researchers have used an AOM device with multiple ultrasonic transducers [23]. These multiple electrodes are fed a complex pattern and launch surface acoustic waves (SAWs) across the device aperture. Diffracted light forms a holographic image. Preliminary results show that this approach may eliminate the need for time multiplexing and consequently scanning mirrors. However, the large

number of electrodes may be prohibitively expensive. Also, the array of SAW electrodes necessitates an additional numerical inversion transformation, making rapid computation difficult.

5.0 Discussion

Real-time 3-D holographic displays are expensive, new, and rare. Although they alone among 3-D display technologies provide extremely realistic imagery, their cost must be justified. Each specific computer graphics application dictates whether holovideo is a necessity or an extravagant expense.

5.1 Applications

I divide interactive computer-graphics applications into two extreme modes of interaction: the "arm's reach" mode, and the "far away" mode. An arm's-reach application involves interacting with scenes in a space directly in front of the user, where the user constantly interacts, moving around it to gain understanding. In this mode, all of the visual depth cues are employed, particularly motion parallax, binocular disparity, convergence, and ocular accommodation. These applications warrant the expense of holovideo and the extreme realism and three-dimensionality of its images: computer-aided design, multi-dimensional data visualization, virtual surgery, teleoperation, training and education (e.g., holographic virtual textbooks on anatomy, molecules, or engines).

At the other extreme, a far-away application involves scenes that are beyond arm's reach and are generally larger. The imagery of such applications - e.g., flight simulation, virtual walk-throughs - make adequate use of the kinetic depth cue, pictorial depth cues, and other depth cues associated with flat display systems. A high-resolution 2-D display may be a more cost-effective solution for far-away applications.

5.2 Present... and Future

Currently there are no off-the-shelf holographic displays. Holographic display technology is in a research stage, analogous to the state of 2-D display technology in the 1920s. What, then, does the future hold? The future promises exactly what holovideo needs: more computing power, higher-bandwidth optical modulation, and improvements in holographic information processing.

Computing power continues to increase. A doubling of computing power at a constant cost - a trend that continues at a rate of every 18 months - effectively doubles the interactive image volume of a holographic

display [5]. Inexpensive computation - around $100 per gigaMAC - is the most crucial enabling technology for practical holovideo, and should be available in 2002.

Although optical modulation has borrowed from existing technologies (e.g., transmissive LCDs, AOMs), new technologies will fuel the development of larger, more practical holovideo displays. Because bandwidth is most important, I use as a figure of merit the number of bits that can be modulated in the latency time of the HVS (typically 20 ms). An AOM can modulate about 16 Mb in this time interval, at a cost of about $2000 (or $120 per Mb), including the associated electronics. The DMD, a new technology for high-end 2-D video projection technology, delivers approximately 100 Mb in 20 ms, for a cost of about $3000, or $30 per Mb. Future mass-production could reduce the cost further. Reflective LCDs are another possible technology. Several researchers create small reflective LCDs directly on a semiconductor chip using VLSI technology [22].

The bandwidths of computation and modulation are likely to increase steadily. Improvements in holographic information processing will likely provide occasional dramatic improvements in both of these areas. Already, holographic bandwidth compression increases fringe computation speed by 3000 times for same-hardware implementation [6]. Standard MPEG algorithms can be used to encode and decode computed fringes [24]. Nonuniformly sampled fringes provide lossless bandwidth compression and promise further advances [25].

User demand may be the one additional key to the development of holovideo2E As other types of 3-D display technologies (e.g., autostereoscopic displays) acquaint users with the advantages of spatial imaging, these users will grow hungry for holovideo, a display technology that can produce truly 3-D images that look as good as - or better than - actual 3-D objects and scenes.

6.0 References

1. Takanori Okoshi, Three-Dimensional Imaging Techniques. New York: Academic Press, 1976.

2. Michael McKenna and David Zeltzer. Three dimensional visual display systems for virtual environments. Presence: Teleoperators and Virtual Environments. vol. 1 #4, pp. 421-458, 1992.

3. David F. McAllister, ed. Stereo Computer Graphics and Other True 3D Technologies. Princeton University Press, Princeton, N.J., 1993.

4. Mark Lucente and Tinsley A. Galyean, Rendering interactive holographic images. Proceedings of SIGGRAPH '95 (Los Angeles, California, August 6-11, 1995). In Computer Graphics Proceedings, Annual Conference Series, 1995, ACM SIGGRAPH, New York, 1995, pp. 387-394.

5. Mark Lucente. Computational holographic bandwidth compression. IBM Systems Journal, vol. 35, # 3&4, pp. 349-365, 1996 Oct.

6. Mark Lucente. Holographic bandwidth compression using spatial subsampling. Optical Engineering, vol. 35, #6, pp. 1529-1537, 1996 June.

7. Mark Lucente. Diffraction-Specific Fringe Computation for Electro-Holography. Ph. D. Thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1994 Sept.

8. Mark Lucente. Interactive computation of holograms using a look-up table. Journal of Electronic Imaging, vol. 2, #1, pp. 28-34, 1993 Jan.

9. Pierre St. Hilaire, Stephen A. Benton and Mark Lucente. Synthetic aperture holography: a novel approach to three dimensional displays. Journal of the Optical Society of America A, vol. 9, #11, pp. 1969 - 1977, Nov. 1992.

10. Pierre St. Hilaire, S. A. Benton, M. Lucente, and P. M. Hubel. Color images with the MIT holographic video display. In SPIE Proceedings #1667 Practical Holography VI, (SPIE, Bellingham, WA, 1992), pp. 73-84

11. Pierre St. Hilaire. Scalable optical architecture for electronic holography. Optical Engineering, vol. 34 #10, pp. 2900-2911, 1995 Oct.

12. Naoki Fukaya, Keiichi Maeno, Koki Sato, Toshio Honda. Improved electroholographic display using liquid crystal devices to shorten the viewing distance with both eye observation. Optical Engineering, vol. 35, #6, pp. 1545-1549, 1996 June.

13. P. Hariharan, Optical Holography: principles, techniques, and applications. Cambridge: Cambridge University Press, 1984.

14. Hiroshi Yoshikawa, Hirokazu Kameyama. Integral holography. In SPIE Proc. #2406 Practical Holography IX (SPIE, Bellingham, WA, 1995), S.A. Benton, editor, pp. 226-234.

15. John A. Watlington, Mark Lucente, Carlton. J. Sparrell, V. M. Bove, Jr., and I. Tamitani. "A hardware architecture for rapid generation of electro-holographic fringe patterns. Practical Holography IX, Proceedings of the SPIE (1995, Bellingham, WA) volume 2406, paper 2406-23, pages 172-183.

16. Dimitri Psaltis, E. G. Paek, and S. S. Venkatesh. Optical image correlation with a binary spatial light modulator. Opt. Eng. vol. 23, #6, pp. 698-704, 1984.

17. F. Mok, J. Diep, H.-K. Liu and D. Psaltis. Real-time computer-generated hologram by means of liquid-crystal television spatial light modulator. Opt. Lett., vol. 11 #11, pp. 748-750, 1986 Nov.

18. S. Fukushima, T. Kurokawa, and M. Ohno. Real-time hologram construction and reconstruction using a high-resolution spatial light modulator. Appl. Phys. Lett., vol. 58 #8, pp. 787-789, 1991 Aug.

19. N. Hashimoto and S. Morokawa. Real-time electroholographic system using liquid crystal television spatial light modulators. J. of Electronic Imaging, vol. 2(2), pp. 93-99, 1993.

20. J. M. Florence and R. O. Gale. Coherent optical correlator using a deformable mirror device spatial light modulator in the Fourier plane. Applied Optics, vol. 27, #11, pp. 2091-2093, 1988 June.

21. Unpublished work by the author, Mark Lucente.

22. Gregory P. Nordin, Jeffrey H. Kulick, Robert G. Lindquist, Michael W. Jones, P. Nasiatka and Stephen T. Kowel. Liquid crystal-on-silicon implementation of the partial pixel three-dimensional display architecture. Applied Optics, vol. 34, #19, pp. 3756-3763, 1995 July.

23. Levent Onural, G. Bozdagi, Abdullah Atalar. New high resolution display device for holographic three dimensional video: principles and simulations. Optical Engineering, vol. 33, #3, pp. 835-844, 1994 March.

24. Hiroshi Yoshikawa, Junji Tamai. Holographic image compression by motion picture coding. In SPIE Proc. #2652 Practical Holography X (SPIE, Bellingham, WA, 1996), S.A. Benton, editor, pp. 2-9.

25. Ravikanth Pappu. Nonuniformly sampled computer-generated holograms. Optical Engineering vol. 35, #6, pp. 1538-1544, 1996 June.

Author:

Mark Lucente
IBM Research Division
Imaging Science and Technology
Thomas J. Watson Research Center
P.O. Box 218, Yorktown Heights, New York 10598 USA
lucente@watson.ibm.com
http://www.research.ibm.com/people/l/lucente
(tel) 914-945-2980
(fax) 914-945-4001

Dr. Lucente is a research scientist in the Imaging Science and Technology group at the IBM T.J. Watson Research Center. He combines knowledge of optics, spatial light modulation, computation, visual perception, and communication systems to develop electro-holography into a practical medium. His work in Visualization Spaces explores how cutting-edge interface and display technologies can be applied to new paradigms of scientific visualization and information management. Dr. Lucente teaches synthetic 3-D imaging at Columbia University in New York where he is Adjunct Assistant Professor of Computer Science. Dr. Lucente worked for six years in the MIT Media Lab Spatial Imaging Group, where he and his colleagues invented 3-D real-time electronic holography. His innovative computational algorithms enabled the world's first interactive holographic imaging in 1990. He earned his college degrees (Ph.D., S.M., S.B.) from the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology. While earning these degrees, Dr. Lucente worked in a variety of fields: 3-D imaging systems, lasers in high-bandwidth communication systems, ultrashort laser pulses, and semiconductor device physics using nonlinear optical phenomena. He is a member of the ACM, OSA, SPIE, Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

7.0 Figures and Figure Captions
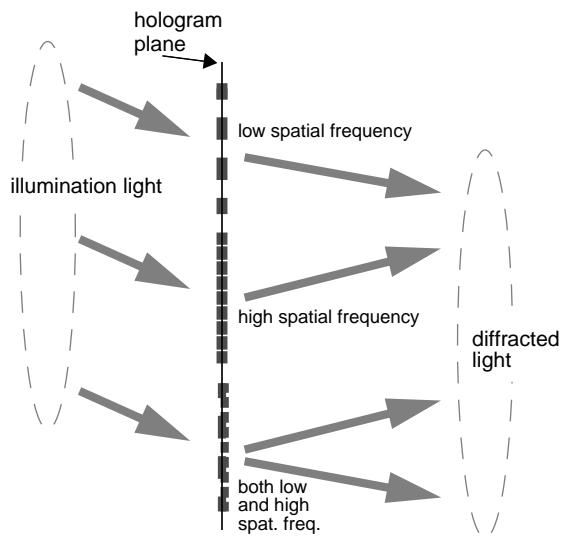
Figure 1: Diffraction of illumination light by holographic fringe patterns. Fringes with higher spatial frequencies cause light to diffract at larger angles. Fringes containing many spatial frequencies diffract light in many directions.
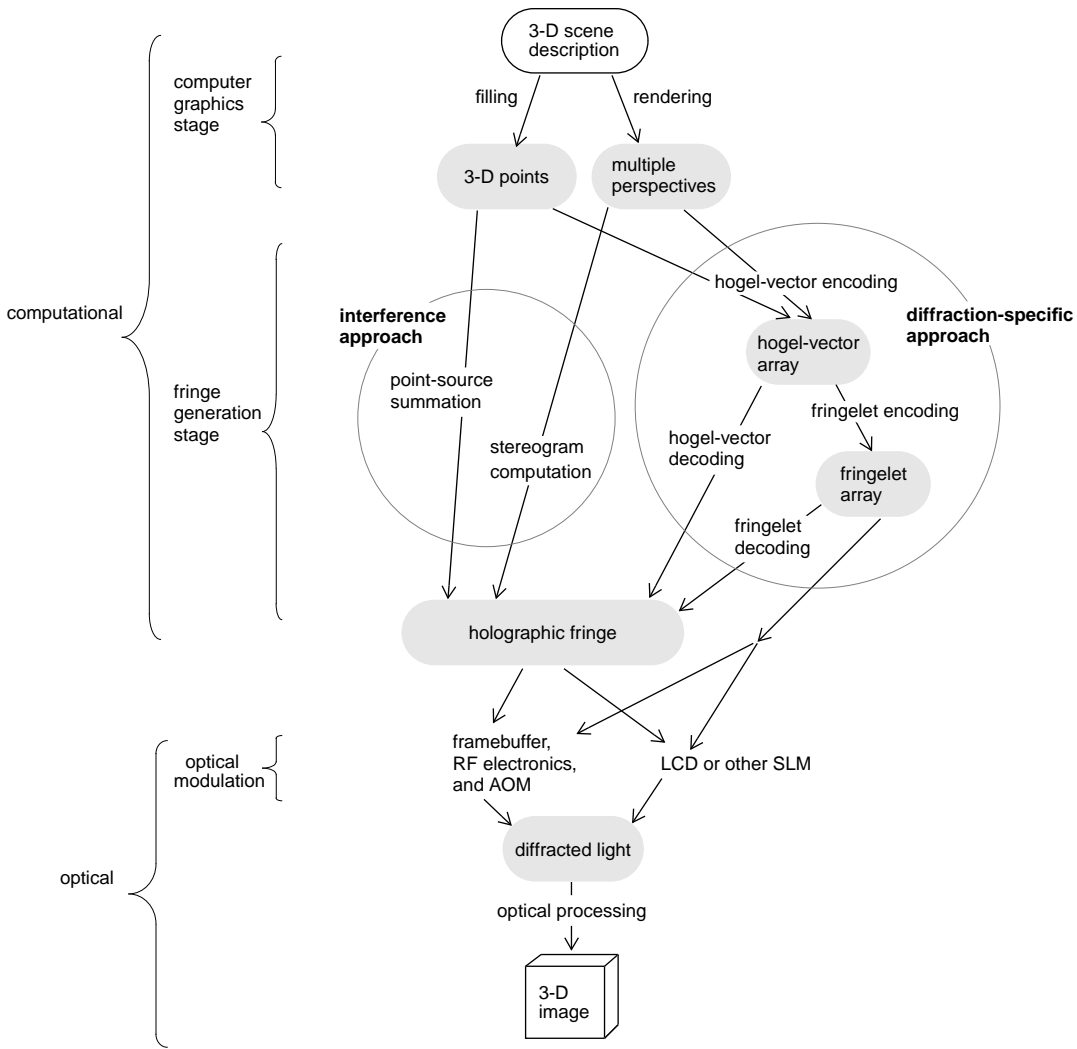
Figure 2: Information flow in interactive 3-D holographic imaging. Each path traces the steps required for a particular method. Computation is generally faster for the methods that are more to the right-hand side.
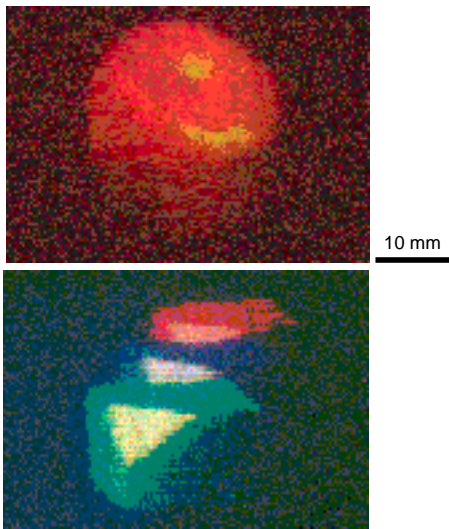
Figure 3: 6-MB holovideo images on the MIT full-color display. Top: reddish apple with multi-color specular highlights, computed using hogel-vector bandwidth compression [5]. Bottom: red, blue, and green cut cubes, computed using stereogram approach [4].
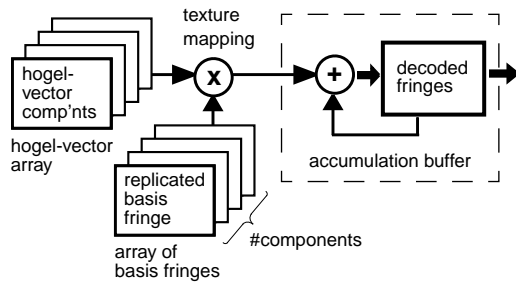
Figure 4: Hogel-vector decoding on the graphics subsystem. The inner product between an array of hogel vectors and the precomputed basis fringes is performed rapidly by exploiting the texture-mapping function and the accumulation buffer.
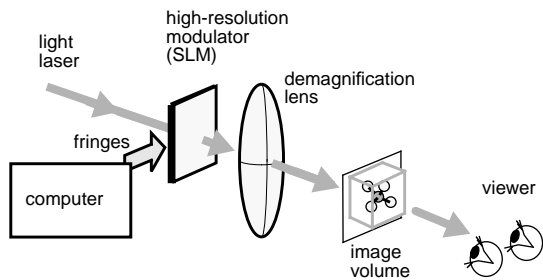
Figure 5: Holographic optical modulation using a typical high-resolution modulator (SLM). A minimum of two million modulation elements is required to produce even a small image the size of a thumb.
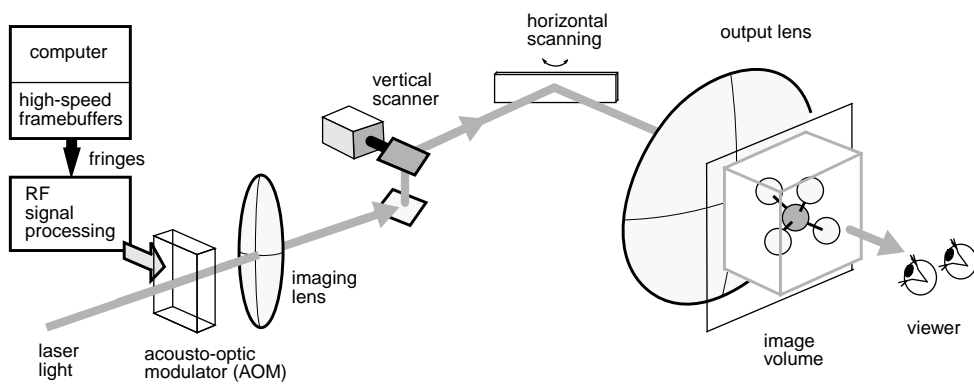
Figure 6: Schematic of the scanned-AOM architecture used in the MIT holovideo displays.

8.0 Glossary

AOM - acousto-optic modulator. A type of high-bandwidth SLM.

CR - compression ratio.

HPO - horizontal-parallax-only. A type of hologram that exhibits horizontal motion parallax horizontally but not vertically.

basis fringe - an elemental fringe precomputed to diffract light in a specific manner.

DMD - deformable micro-mirror device. A micromechanical SLM.

fringe - the holographic pattern that is either recorded optically or generated computationally and used to diffract light to form an image.

HVS - human visual system.

hogel - holographic element. A small piece of hologram that has homogeneous diffraction properties.

LCD - liquid crystal display. An electro-optic SLM.

MB - 1048576 bytes.

MAC - multiplication accumulation. A numerical calculation consisting of one multiplication and one addition.

SLM - spatial light modulator. A device that modulates a beam of light.