Spatial Interaction with Haptic Holograms

Wendy Plesniak and Ravikanth Pappu
Spatial Imaging Group, MIT Media Lab
20 Ames Street E15-416
Cambridge, MA 02139. USA.
{wjp, pappu}@media.mit.edu

Abstract

We present a comprehensive report of our work in the design and implementation of systems that combine holography with haptic feedback. Two separate ***holo-haptic*** *systems are described. First, we describe a static system wherein a user can see a free-standing* ***static holographic image*** *and inspect its shape using a* ***force-feedback*** *stylus. In the second system, visual display is provided by a* ***dynamic holographic video*** *system, and the user can modify the geometry of the image in near real-time. We also suggest a preliminary set of guidelines for presenting* ***physically believable*** *multimodal simulations in coincident visual-haptic workspaces.*

Keywords: *multimodal interaction, virtual and augmented reality, holography, haptics.*

1. Introduction

The last decade has witnessed a great deal of interest in building systems with multimodal input and output capabilities. The availability of large computational bandwidth coupled with innovative sensing and display technologies has inspired a varied and interesting array of spatial interactive workspaces [1][2][3][4][5]. At the same time, there is growing experimental evidence that the availability of binocular visual depth cues, spatial audio, kinesthetic and tactile feedback improve performance of some spatial tasks [6][7][8]. These studies underscore the need for multimodal workspaces with spatial cueing, and motivate research on bringing new sensing and display technologies into the mainstream.

In the Spatial Imaging Group at the MIT Media Lab, we build experimental spatial displays. We believe that by engaging binocular vision and motion parallax, these displays offer a powerful way to disambiguate spatial information in a scene and help a viewer to better understand the shapes and layout of displayed objects. Although we investigate several different kinds of 3D imaging techniques, holography is certainly the most prominent among them.

In both the systems described in this paper, we employ holography to generate visual spatial images and combine them with co-located force display. First, we describe a system wherein a user can see a free-standing, static holographic image and inspect it using a force-feedback device. In the second system, spatial visual imagery is provided by a dynamic holographic video system, and, using the same force-feedback device, a user can inspect and *modify* the geometry of the image in near real-time. Both systems provide *coincident workspaces*, wherein the haptic and visual feedback originate from the same spatio-temporal location.

The structure of this paper is as follows. First, in the context of coincident workspaces, we compare holographic displays to other techniques for displaying 3D images. Then, for each of our holo-haptic systems, we present its architecture, operation, and discuss system performance and related issues. Finally, we suggest a preliminary set of guidelines for presenting physically believable multimodal simulations in coincident visual-haptic workspaces.

2.0 Why holographic displays?

Why should we use holographic displays in coincident workspaces? In this section, we present a concise overview of three popular alternative techniques for 3D display, and compare them to holography.

2.1 Stereoscopic systems with beamsplitters

A historically popular and frequently used technique employs a half-silvered mirror to combine two-view or head-tracked stereo images with a view of the hand and physical objects in the manual workspace. In these systems, the stereo graphics (or video) can neither occlude the

real world, nor can they be occluded by it (unless hands or objects are tracked, and corresponding occlusion computations are performed.) The result is a "ghost-like" display in which simulated and real objects sometimes appear to be transparent. In such a configuration, volumes of simulated and real objects can freely interpenetrate; when co-located, their appearance is strange, and their absolute depths may be inaccurately perceived [9].

When stereoscopic video or computer graphics are used to supply the display, depicted objects can be realistically rendered and displayed at high resolution. 3D scenes thus generated can be replete with pictorial cues to depth and can be rendered with convincing shading, texture and reflections. However, pixellated stereoscopic displays have an inherent depth quantization associated with them; for a CRT with 0.25mm pixel pitch, and a viewer with an interpupillary distance of 65mm located 500mm from the display plane, the depth quantization is on the order of 2mm. Since human stereo acuity is on the order of 1mm at this distance, these displays cannot yet provide adequate depth resolution.

This kind of system also fixes a scene's depth of field during rendering or scene optical capture, often without considering an appropriate depth of focus for viewers' eyes. Resulting stereo images often have an abnormally large depth of field which may account for the distracting nature of unfuseable images in stereo displays (while diplopic vision goes virtually unnoticed during normal binocular viewing).

If head-tracking is employed to provide motion parallax, scene jitter from tracker noise and lag between fast head movement and scene update can be problems. Additionally, some sort of viewing apparatus must be worn, such as LCD shutter glasses, to multiplex the correct stereo views to left and right eyes (and often to provide head-tracking information to the system). Though these glasses are not particularly uncomfortable, autostereoscopic viewing is generally preferred, especially for viewers who already wear glasses.

Inherent in stereoscopic and other astigmatic displays is an accommodation-convergence mismatch—a functional disengagement of several systems of the visual system which normally behave in cooperation. Stereoscopic systems require a viewer to accommodate to (or near) the display plane to see the imagery in focus. Yet, 3D imagery is likely designed to also appear at different depths than that of the display surface. To fuse this imagery, a viewer must converge to a location different than the display surface while keeping the image in focus. This mismatch does not necessarily frustrate the perception of the three-dimensionality of images; the link between accommodation and convergence is plastic and many displays disrupt it with no serious perceptual effects. However, the extent to which accommodation-convergence mismatch may be responsible for errors in perception and visually-guided movement is undetermined.

In coincident visuo-manual applications using astigmatic displays, there is also an accommodation mismatch between simulated and physical objects. For instance, when an operator moves a finger or a hand-held tool into zero-disparity alignment with a feature on the simulated object, conflicting depth cues may be reported by accommodation.

2.2 Volumetric displays

Volumetric displays work by sweeping or scanning out 3D space with points or planes of light. In these displays, solid objects are approximated by a spatial arrangement of image slices or by many points of light spatially assembled to indicate object structure. There are several types of volumetric displays available: slice-stacking displays, like the varifocal mirror [10], rotating plane/helix displays, such as the Texas Instruments *OmniView*; and systems which actually emit photons from within the display volume.

These displays have some advantages: and there is usually no mismatch between accommodation and convergence since all "object points" have a true spatial location to which eyes can freely converge and focus, and they typically have a wide field of view. However, the low density with which many such systems display object points causes depth quantization, and since the light emitting points in these systems are isotropic radiators, view-dependent shading is not possible. Worse still, inter- and intra-object occlusion relationships are not displayed (unless viewer position is tracked and corresponding backface cull and occlusion computations are performed on the object point database). Since occlusion is almost always the most powerful cue to depth in a scene, images appear translucent without it, and the impression of three-dimensionality can be seriously compromised.

Finally, in coincident visuo-manual applications, these displays are mostly entirely inappropriate —most cannot physically admit the hand into their display volume.

2.3 Re-imaging systems

Re-imaging display devices use optical systems to com-

bine or condition images and relay them into a viewer's space. For instance, Dimensional Media's *High Definition Volumetric Display* and SEGA's *Time Traveller* arcade game relay images of real 3D models and 2D computer graphics (flat or pre-distorted) to a viewer. This class of systems employs optical components such as parabolic mirrors, lenses, and beamsplitters to re-image *already-existing* 3D objects and/or 2D display screens. As a result, interacting with the optical output in a way that modifies the true shape of the displayed object is not possible [11].

Some of these systems, like DMA's *HoloGlobe*, are capable of displaying a large free-standing image with a wide field of view. In addition, a viewer can see the image under ambient lighting conditions and without wearing viewing aids. Accommodation and convergence probably behave the same as they do during normal binocular viewing. Additionally, since these systems re-image existing physical 3D models, the visual realism of the display can be quite striking.

However, in a coincident visuo-manual workspace, it is possible for the interacting hand to literally block image projection. If the spatial image being occluded by the hand is located nearer to the viewer's eyes than the hand, the available cues to depth become strongly contradictious æ binocular cues and motion parallax report accurate depth relationships while occlusion reports the opposite depth ordering. The projective geometries of re-imaging displays make them prone to this particular problem.

2.4 Holographic displays

Holographic displays permit this problem to arise as well. If a person's hand is interposed between object points (nearer the viewer's eye) and the farther hologram plane, image reconstruction is blocked and the hand appears to occlude a nearer object. While this cue conflict is disturbing, it occurs only during certain configurations of the hand and object; this is less distracting than the ever-present hand/image composite displayed by stereoscopic systems with a half-silvered-mirror. However, holographic displays, which are actively being researched and developed, have increasingly many advantages to offer.

A holographic stereogram is a *discretized* hologram [12] which optically projects a series of 2D perspective views of a scene into the display's viewzone. There, a person is provided autostereoscopic viewing of the scene, and scene parallax changes (without lag) in concert with the viewer's head motion. Holographic stereograms offer flexible image content, giving them an advantage over re-imaging displays; 2D component images can be realistically rendered using computer graphics or optically captured with a scanning camera and computationally predistorted when necessary [13]. Thus, holographic stereograms can be replete with pictorial cues to depth in addition to binocular and motion parallax, and *full parallax* holograms (in principle) allow a person to freely converge and accommodate to any depth in the image volume. When computer graphic or other digitized component images are used to generate the display, however, holographic stereograms are subject to the same depth quantization of any system using a pixellated display screen.

From a technological standpoint, holographic stereograms are becoming faster and less expensive to produce [14]. The availability of new recording materials and processing techniques have improved diffraction efficiency and signal-to-noise ratio, yielding brighter, cleaner images. New display formats that incorporate their own illumination (edge-illuminated) are becoming compact and portable [15]. And finally, in recent years, research on electronic holography has yielded systems which provide full color, moving holographic images [16].

In order to project a visual spatial image into the viewer's manipulatory space, display flexible image content, and provide as many cues to depth as possible, we chose to incorporate holographic displays into our spatial visuo-haptic systems. Our first experiment involved the combination of static edge-illuminated holograms and coincident force display.

3.0 Static haptic holography

The combination of haptics and holography was first investigated by researchers at De Montfort University for an object inspection task [17]. In this work, visual display was provided by a *reflection* transfer hologram which presented an aerial image of a control valve. A Computer Controlled Tactile Glove (CCTG) provided coincident haptic display of the same data. Early informal experiments in combining reflection transfer holograms with force-feedback were also performed by researchers at the MIT Media Laboratory's Spatial Imaging Group. Reflection holograms require front overhead illumination for image reconstruction; thus, in either of these holo-haptic efforts, the interacting hand could literally block the reflection hologram's illumination and prevent image reconstruction.

This problem was addressed in our laboratory by

employing full-parallax edge-illuminated holograms in combination with the Phantom™ force-feedback interface for the inspection of static 3D models [18]. In this work, the systems for optically recording holograms and producing the haptics simulations are separate, but both take the same 3D geometry description as input. The hologram-production pipeline also requires specification of camera geometry, scene lighting information and all visual material properties for rendering. The haptics pipeline requires tactual and bulk material property specification to accompany the geometric description.

3.1 Implementation

3.1.1 Static hologram modeling and printing. We have developed a computer graphics rendering server (*HoloServe*) and client application for both visual scene design (*HoloBuild*) and holographic stereogram printing (*HoloPrint*).

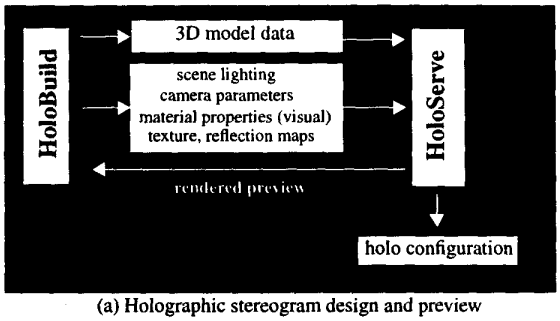
HoloBuild makes it possible to:

- import 3D model data,
- interactively light the scene,
- tailor the computer graphics view and camera to match hologram geometry,
- change model rendering parameters and
- request that a specific perspective view be rendered by HoloServe.

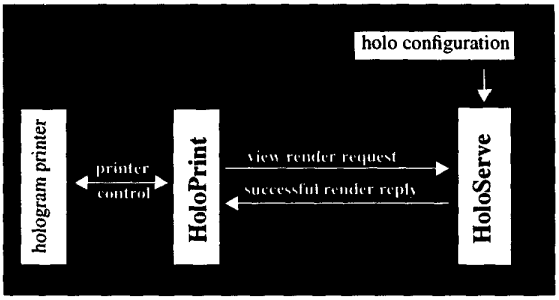
HoloServe uses SGI's Graphics Library and takes advantage of available rendering hardware; individual perspective views of complicated computer graphic models can be generated in a few seconds.

All final model and rendering parameters designed using HoloBuild are saved in a configuration file. This file is read directly by HoloServe for rendering component images at hologram printing time. During the hologram printing process, the HoloPrint application requests that a specific perspective view be rendered, and upon notification of rendering completion, sends appropriate exposure and frame-advance control sequences to the hologram printer. In this manner, the newly rendered frame is recorded and the printer is readied for the next exposure (Figure 1b). The combination of HoloServe and HoloPrint marked our first "render-on-demand" holographic printer, obviating the need for storing large numbers of pre-rendered frames on disk for subsequent hologram recording.

Both HoloBuild and HoloServe were designed to accommodate a wide variety of hologram viewing geometries, and both one-step and multiple-step printing processes. For any given optical printing setup, the specific geometries of the hologram printer can be specified in HoloBuild during the design process, recorded in the configuration file, and used later during hologram printing.



(a) Holographic stereogram design and preview



(b) Holographic stereogram printing pipeline

Figure 1. Design and Preview, and "Render-on demand" hologram printing pipelines

HoloServe, as well as HoloBuild, were implemented on an SGI Onyx Reality Engine, and HoloPrint can be run on any unix machine. These design-preview rendering and printing processes communicate over ethernet using remote procedure calls (rpc).

3.1.2 Haptic Modeling. The haptic simulation, (*HoloFeel*) uses the same geometrical description as HoloServe to render a force image of the scene in space, which we co-locate with the projected holographic image. From a client scene design tool, HoloFeel receives the particular scene geometry and haptic modeling parameters that correspond to a given holographic image (Figure 2). Then, to display the simulation computed with HoloFeel, we use the Phantom haptic interface, a three degree-of-freedom (d.o.f.) mechanical linkage with a three d.o.f. passive gimbal that supports a simple thimble or stylus used by the hand. Six encoders on the device provide positional information, and three servo motors provide force display in a workspace of 290 x 400 x 560mm^3. HoloFeel is implemented on a Pentium PC and runs with an average servo

rate of 2KHz.

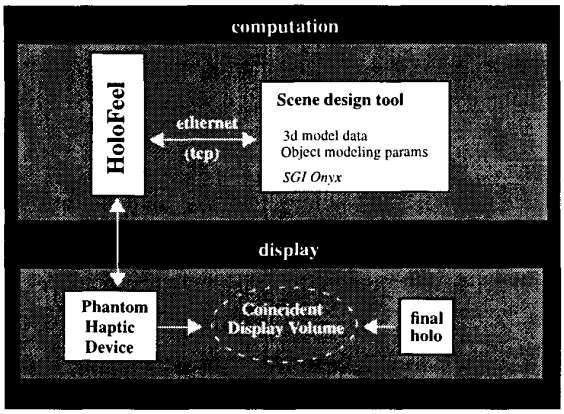


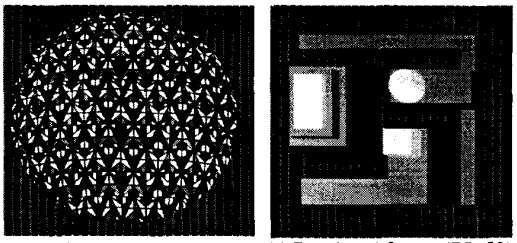
Figure 2. Static holo-haptic system architecture

3.2 Results

This suite of applications was used to produce two two-optical-step edge-illuminated holographic stereograms used in conjunction with haptic simulations. The simpler hologram displayed a hemisphere affixed to a vertically oriented plane. The plane was both visually and haptically textured with a vertical grating; the hemisphere itself was visually texture mapped but had no haptic surface relief. Haptic modeling included static and dynamic friction, and object damping and compliance. A rendered component frame and the final hologram are shown in Figure 3(a) and Figure 4(a).

The full-parallax master hologram (first optical step) permitted a broad angular range of head motion of approximately 50 degrees horizontally and 30 degrees vertically. The master contained 15,000 exposures of pseudoscopically-rendered [13] frames. The final edge-illuminated hologram was printed in an additional optical transfer step. Mastering and transferring steps [14] used a recording wavelength of 528nm.

The total depth of the final hologram was approximately 40mm, all in front of the image plane. Image plane width and height were each 100mm. The hologram was illuminated with an LED centered at 520nm, which yields a bright image with slight spectral blurring. The multi-modal scene presented was intended to have very few formal features; this tangible hologram provided a simple example with which to examine perceptual tolerances for spatial misregistration and mismatches in curvature of the visual and haptic models (*e.g.* by replacing the haptic hemisphere with a hemi-ovoid).



(a) Rendered frame (40, 50). (b) Rendered frame (75, 50).

Figure 3. Rendered component images

The second and slightly more complex holo-haptic example used an arrangement of blocks forming a maze, which is oriented against a vertical back plane. The blocks varied in size and spacing, and the channels formed between them are narrow. The back plane was visually and haptically texture mapped with a vertical grating. A rendered component frame and the final hologram are shown in Figure 3(b) and Figure 4(b).

The full-parallax master provided a much smaller horizontal range for head motion in the final hologram æ approximately 35 degrees. The master was comprised of 6700 exposures of pseudoscopically-rendered frames [13] and the final hologram was produced in an additional optical transfer step. As in the previous hologram, the mastering and transferring steps used a recording wavelength of 528nm.

The total depth of the final hologram was approximately 35mm, and the entire model reconstructed in front of the hologram plane. Image plane width and height were each 100mm. The final hologram was illuminated with an LED centered at 520nm. The model presented contains more image features (edges) than did the simple hemisphere hologram; the maze hologram offered us a chance to examine maze-tracing performance in this coincident and an offset (using stereo graphics) visuo-haptic workspace configuration.

In addition to the robust and compact design of edge-illuminated displays, the principal benefit of using this hologram format in concert with haptic applications lies in the steep-angle lighting it incorporates; in the final viewing configuration, the hand and haptic apparatus do not block the illumination source as they interact with the image. However, while haptic inspection of the holographic image is possible using this hologram format, we still encounter some discord between the hologram and

visible co-located hand and haptic apparatus. This slightly more complex volume of information provided greater opportunity for us to examine some of the sensory conflicts which occur in a coincident holo-haptic workspace.

(a) haptic inspection of hemisphere hologram

(b) haptic inspection of block hologram

Figure 4. Holograms combined with force image

3.3 Sensory Conflicts

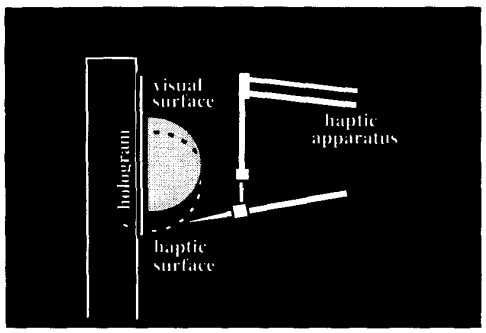
As we readily observe in our everyday interactions, harmonious multisensory stimulation usually gives rise to correct perception of objects and events. The broad body of work on multi-sensory interaction indicates that some disparity between visual and haptic information can distort the overall percept while still being tolerated. The ability of sensorimotor systems to adapt to discordant sensory input permits us to perform well even in the presence of distortion, so long as sensory feedback is available. This fact is extremely useful in offset visuo-haptic workspace configurations, wherein the tracked hand or device position is represented as a graphical element on the visual display and the user never actually visually observes her hand. In such workspace configurations, slight spatial misregistrations, or changes in scale between the visual and haptic display can be virtually unnoticeable. Yet too much intermodality disparity can cause the visual and haptic cues to be perceived as entirely separate events, and may be quite confusing or annoying.

Tolerances are lower still when visual and haptic workspaces are superimposed. In our coincident workspace format, we observed several conflicts between what is seen and what is felt; these intra- and intersensory conflicts are described in turn below.

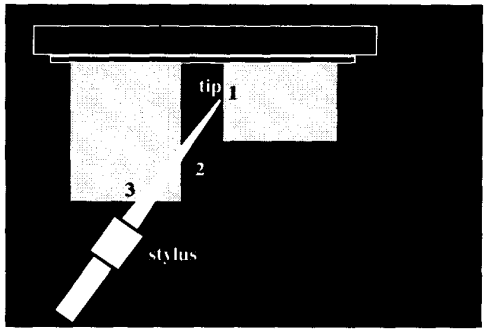
3.3.1 Spatial misregistration. When exploring a surface with the Phantom and visually monitoring the device, simultaneous visual and haptic cues to the surface location are available. When we *feel* contact, the visible location of the stylus tip is perceived to be co-located with the haptic surface. During contact, if the holographic surface and the haptic surface are not precisely aligned, the misregistration is strikingly obvious to vision. These conflicting visual cues erode the impression of sensing a single object. Instead, the impression of two separate representations is evident. This condition is shown in Figure 5(a).

3.3.2 Occlusion violations. As mentioned earlier, occlusion is perhaps the most powerful cue to layout in a scene. When we see the image of an object being blocked by the image of another, we understand the occluded object to be farther from our eye than the occluding one. In our holo-haptic systems, it is possible to position the haptic apparatus between hologram and image and actually block its reconstruction; in an observer's view of the scene, occlusion relationships contradict other depth cues reporting true scene layout as shown in Figure 5(d). Even in the presence of correct depth accounting from stereopsis and motion parallax, Perception appears to favor the depth ordering reported by occlusion relationships.

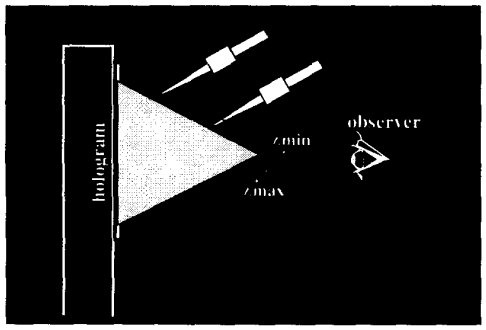
3.3.3 Contact with multiple surface points. Obviously, holograms present spatial images which cannot by themselves exhibit a restoring force when penetrated by an object. With no haptic simulation running to detect collisions with model surfaces and to display contact forces, the haptic apparatus is free to pass through the holographic image undeterred. Our haptic simulation can prevent a single point on the stylus from penetrating the model, but current device limitations preclude emulation of the kind of multi-point contact that occurs in the physical world.



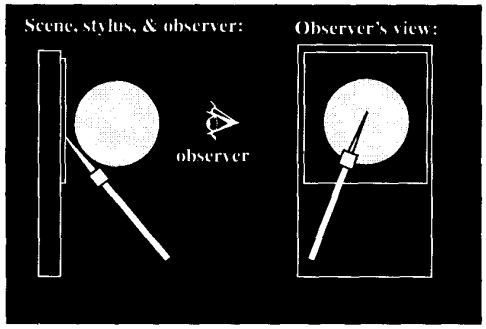
(a) Spatial misregistration results in visual disparity between force and optical images.



(b) Restoring force is displayed due to surface penetration at **1**, while the visible surface penetration from **2** to **3** is not haptically reported.



(c) Broad spectrum illumination of hologram causes depth-dependent optical image blur.



(d) In the *true* spatial layout, the stylus is positioned behind the holographic image of the sphere. In the *apparent* depth ordering, the stylus is seen occluding (in front of) the sphere.

Figure 5. Conflicting sensory cues present in holo-haptic displays

During each haptic control loop cycle, the simulation checks for a surface collision all along the stylus probe; even if it finds many, it can only compute and display forces for one. If a model surface has been penetrated by the stylus tip, it is assumed the viewer's primary attention is focused there, and forces due to this collision are computed and displayed. However, if not the tip, but other points along the probe have penetrated the model, then the collision closest to the tip is used for computation and display.

The situation permits another kind of occlusion violation to occur as shown in Figure 5(b). While the stylus tip is seen and felt in contact with some geometry, the stylus may be rotated around its tip and swept through proximal holographic image volume. Parts of the user's hand may also penetrate the image. Seeing such physical objects and holographic image coexist in the same physical volume presents a confusing impression of depth and object solidity in the scene.

3.3.4 Optical-haptic surface property mismatch. An artifact of a hologram's diffractive properties is the chromatic blurring that occurs with broad spectrum illumination. In the transmission edge-illuminated holograms used in this work, the holographic image plays out high and farther from the hologram in wavelengths shorter than the recording wavelength, and lower and closer in longer ones. If the illumination source used in hologram reconstruction is not monochromatic, spectral blur will be evident in the final image. Image elements close to the hologram plane will be quite clear, but those farther from the hologram plane will exhibit blur in accordance with source bandwidth.

Since a viewer generally expects scene elements closer to the eye to be more keenly resolvable, the blurry image elements near the viewer challenge the impression of image solidity. This condition, shown in Figure 5(c), is recognized as problematic on its own, but adding coincident haptic display causes further difficulty. Usually an object's bulk material properties (*e.g.* stiffness) remain uniform throughout the display volume. If the haptic and

visual output are precisely in register, then near the hologram plane the stylus will be exactly coincident with an imaged surface during contact. However, far from the image plane, the stylus will visually penetrate the blurry image of the surface by a substantial distance before contact is felt. As mentioned earlier, misregistration between the image surface and stylus tip during contact, especially when close to the viewer's eye, can diminish the simulation quality. In addition, visual and haptic information presented is conflicting; by visual report, the surface qualities change substantially with depth though their haptic quality remains the same.

In addition to the occurrence of these cue conflicts, the image quality was somewhat compromised, mostly due to non-uniform illumination. This problem is not inherent to the display format; better illumination can certainly be devised. Nonetheless, these full-parallax haptic-holgrams presented a compelling way to haptically inspect simulated spatial objects using a compact table-top display. For instance, when a person taps the edge of a block, or visually and haptically locates a trough in an object's texture with the stylus tip and follows it through the image, the impression of a single multimodal representation is quite strong. The principal disadvantage to these displays (and no small one at that) is that they are static; they have utility for inspecting three dimensional shapes, but do not permit interaction with or modification of the data presented.

4. Dynamic haptic holography

The logical extension of using the Phantom to inspect a static holographic image is to allow an operator to modify the geometry of the image using the Phantom. This requires (ideally) a real-time, dynamic holographic display. For the past several years, one of the research projects in our group has been to build a real-time holographic video display (*holovideo*). Combining haptic simulation and force feedback with holovideo permits us to render *dynamic* scenes in a coincident workspace.

4.1 System architecture

Two separate modules comprise the computation which feeds the displays; a *haptics module* that performs force modeling, and the *holovideo module* which pre-computes holograms and drives rapid local holographic display updates based on changes to the model. The haptics and hologram modules are organized by the *Workspace Resource Manager* (WRM) which is notified of geometry changes imparted to the haptic model by the user, and requests hologram updates to local regions of the visual display where changes have occurred. From the point of view of a user, who is holding the stylus and pressing it into the holographic image, a single multimodal representation of the simulation can be seen and felt changing in response to the applied force. The system architecture is shown below in Figure 6.

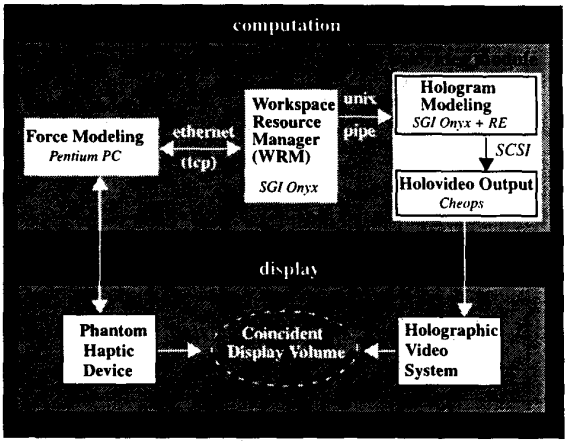


Figure 6. Dynamic holo-haptic system architecture

4.2 Haptic Modeling

The multimodal image displayed represents a spinning surface of revolution (initially a cylindrical "stock") which can be interactively "lathed" by user. The haptic cylinder, initially and in subsequent stages of carving, is represented as a surface of revolution with two caps. It has a mass of 1gm, an algorithmically defined vertical grating as surface texture, static and dynamic frictional properties, stiff spring bulk resistance, and rotates about its axis at one revolution per second. The cylinder model straddles a static haptic plane (which spatially corresponds with the physical output plane of the holovideo optical system); the haptic plane is modeled with the same bulk and frictional properties as the cylinder. Currently, the haptics simulation is implemented on a Pentium PC with an average servo rate of 1.5KHz.

The radius profile of the surface of revolution is represented as a cubic B-spline curve with 28 control points, all of which are initially set to the same radius value (25mm) to let us begin with a cylinder. The curve evaluated between the middle 21 points defines the profile of the cylinder body; the remaining top three and bottom four points lie beyond the actual extent of the cylinder, and serve to "lock" the shape at its top and bottom, respectively. Control points are modified as force is exerted on the shape at height h, between control points P_i and P_{i+1}.

A new radius for the surface of revolution at this height can be computed by evaluating the nonuniform rational B-spline formulation.

The cylinder can be felt spinning beneath the user's touch, and when pressed with enough force (*i.e.*, when the surface has been penetrated by some threshold distance D) the surface deforms. A very simple method for surface deformation is used: the two control points straddling the penetration location are displace toward the central cylinder axis by a fraction of the penetration distance, and this changes the radius profile of the surface of revolution.

The upper control point is displaced by tkD, and the lower by $(1\text{-}t)kD$, with t being the normalized distance between the contact point and the lower control point, used in the B-spline formulation. The closer control point is displaced by a greater distance. If contact occurs directly on a control point, then that point alone is displaced by kD. Thus, control point displacement modifies the circumference of the cylinder at height h, as force is interactively applied.

The parameters k and D can be adjusted to make carving the rotating cylinder require more or less force. A minimum radius of 15mm is enforced, so that once the surface has deformed this much, the control points update no further. The control point density, 4.17 points/cm, was experimentally determined to be high enough to accommodate local model changes, yet sparse enough to avoid unstable deep notching of the haptic surface.

4.3 Holographic Video Modeling

We employ the second generation of holovideo in this work [16]. This system is capable of displaying monochromatic, horizontal-parallax-only (HPO) images in a volume of 150 x 57.5 x 150mm^3, and the viewing angle is 30°. The 3D image produced by holovideo supports the most important depth cues: stereopsis, motion parallax, occlusion, and many pictorial and physiological cues to depth.

For the present purpose, we may consider holovideo to be a black box which accepts two inputs: a *computer-generated hologram* (CGH) and light [19][20]. The output of the black box is a 3D holographic image whose visual and geometrical characteristics depend on how the CGH was computed. Each CGH contains an enormous amount of data — 36 megasamples (at 1 byte per sample) apportioned into 144 hololines of 256 kilosamples each. The CGH is made available to the display via a framebuffer. Because holovideo has a non-standard display format, an image-processing system developed at the MIT Media Lab, *Cheops*, was extended to support it. Cheops has three different module types: processor, input/memory, and output, and an optional memory module provides up to 0.5 Gbytes local to the system. These modules are interconnected by two linear buses. One of these buses, the Nile bus, is capable of sustained high bandwidth (>100 Mbyte/sec.) transfer of samples and the second, the Global bus, is capable of 32 Mbyte/sec. transfer [21].

4.4 Implementation

4.4.1 Haptics Module. The Workspace Resource Manager (WRM) initializes its own model of the surface of revolution, which starts as a cylinder of desired height and radius. It then initiates the haptic simulation by making client calls to the haptics module on the Pentium PC. These calls request creation of a haptic cylinder of the same height and radius at a desired location. The haptics module commences physical simulation of this spinning cylinder, and computes collisions of the Phantom tip with the computational model. Based on these collisions, forces are computed and displayed to the operator's hand, and any resulting shape modifications are reflected in the model update.

Changes in the cylinder's underlying B-spline representation are automatically communicated from the haptics module to the WRM approximately 30 times per second. The information sent contains the location where change begins on the curve (the number of the bottom-most control point), and values of the six affected control points, ordered from bottom to top. It is assumed that model changes occur reasonably slowly, so that no more than six control points are updated within 33ms. Since computing a deformation means updating at most two control points surrounding the point of contact, our communication rate means that we can only guarantee reporting accurate model changes from contact in a region 6.9mm high within an update interval. Though this assumption usually puts us within the realm of normal interaction speed, eventually, communicating a variable number of control points to reflect the precise region of change would be more robust, and future work will implement this change.

4.4.2 Workspace Resource Manager. Once the WRM receives the message, the changed control points are used to update its own representation of the radius profile. The

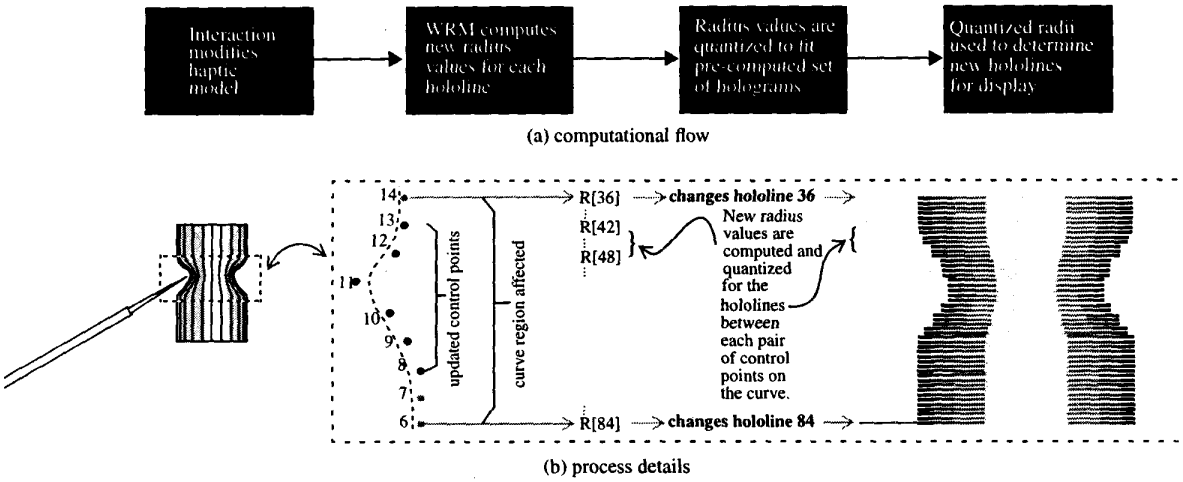


Figure 7. Method of propagating haptic model changes to holographic display

WRM determines which lines of the holovideo display will be affected by the updated region of the curve. Since the final holographic image will span 120 lines of the display, we maintain a state vector, R, with 120 elements whose values represent the exact radii of the surface of revolution at corresponding display lines. A set of six holovideo display lines correspond to the space between any two adjacent control points in the WRM's model.

If as many as six control points have changed, it is necessary to recompute radii for the 48 display lines spanning *eight* control points, between which the curve will have been affected (Figure 7). These new radius values are reflected in the state vector **R**. In the current implementation, the WRM's model can also be rendered to a graphics display using SGI's Graphics Library for debugging purposes, and to provide a means for remotely monitoring a user's performance.

Because it is not yet possible to compute 36 Mbyte holograms in real time [19], we decided to pre-compute five cylinder holograms for use in updating the display, as explained shortly. Each hologram displays a cylinder with a different radius, the initial cylinder, and four progressively smaller ones, r_{cyl} (mm) = {25.0, 22.5, 20.0, 17.5, 15.0}, ending with the minimum-radius cylinder. All holographic cylinders are 47.9mm high. These holograms, from largest to smallest radius, are loaded sequentially into the Cheops memory module. At system start-up, the cylinder with the largest radius is displayed. As the initial haptic cylinder is carved, a visual approximation to the resulting surface of revolution is assembled on the display by loading the appropriate lines from each of these five separate holograms.

First we must determine *how many* and *which* lines we should change on the holovideo display. The number of display lines that require updating will vary, depending on exactly which model control points are displaced. In regions near the top or bottom of the carved shape, a smaller region of the curve contributes to the visible extent of the shape, so fewer display lines will require change. The new radius values in **R** corresponding to changed display lines are quantized to match one of the set of five holographic cylinder radii, and each is assigned a radius code based on its quantized value as shown below:

radius (mm)	25.0	22.5	20.0	17.5	15.0
code	5	4	4	2	1

A message, which contains the number of the hololine marking the start of the update region, the number of lines that need to be updated, and the radius codes of each new line, is sent to the holovideo output module on Cheops. In order to minimize the display update time, we are currently updating a maximum of 32 hololines per cycle, representing only the display lines between the original six control points sent by the haptics module.

4.4.3 Holovideo Indexing. Upon receiving the update message, the holovideo output module must instruct Cheops to collect the appropriate hololines and dispatch them to the display. This is accomplished by indexing into the memory module with the radius code to determine the correct cylinder to display, and then writing the corresponding hololine to the output card (Figure 8). The final holographic image is assembled using hololines

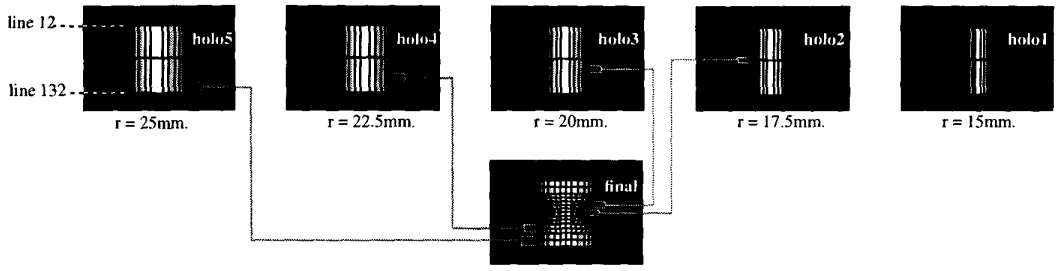


Figure 8. Method of assembling final holographic image from pre-computed hologram set

from the five individual holograms. It must be noted that this method of hologram assembly is valid only for HPO holograms; for full-parallax holograms, the entire hologram would have to be recomputed. In the absence of the computation and communication bandwidth necessary to update fully-computed holograms in real-time, pre-computed hologram indexing enables rapid, local updating.

4.5 Results

When an operator carves the holographic surface of revolution with the Phantom, the hologram image changes due to force apparently applied by the tip of the stylus. The resulting shape can be explored by moving the stylus tip around the surface without exerting too much force. Physical objects in the workspace may also be explored, so that both physical and simulated forces can be displayed to the operator alternatively in the same workspace. When the operator maintains the correct viewing distance for holovideo, the perception of a single multimodal stimulus is quite convincing. Images of an operator interacting with the image are shown in Figure 9. Different stages of carving are also shown in the lower part of the figure.

4.5.1 System Lag. A compelling multimodal representation depends heavily on minimizing, to imperceptible levels, the time lag between the operator effecting changes in the haptic model and the result of that change appearing on the visual display [22]. A reasonable visual update rate (20+ frames per second) is not currently possible on holovideo, principally due to the speed at which we can communicate with and update the display. The effect of the resulting system lag, on the order of 0.5 sec., is that an operator can *see* the stylus tip penetrating into the holographic surface before the surface is apparently subtracted away. Higher bandwidth spatial light modulators, efficient data compression techniques, improvements in computation speed, and higher bandwidth data pipelines will all help to alleviate this problem in future generations of the holovideo system.

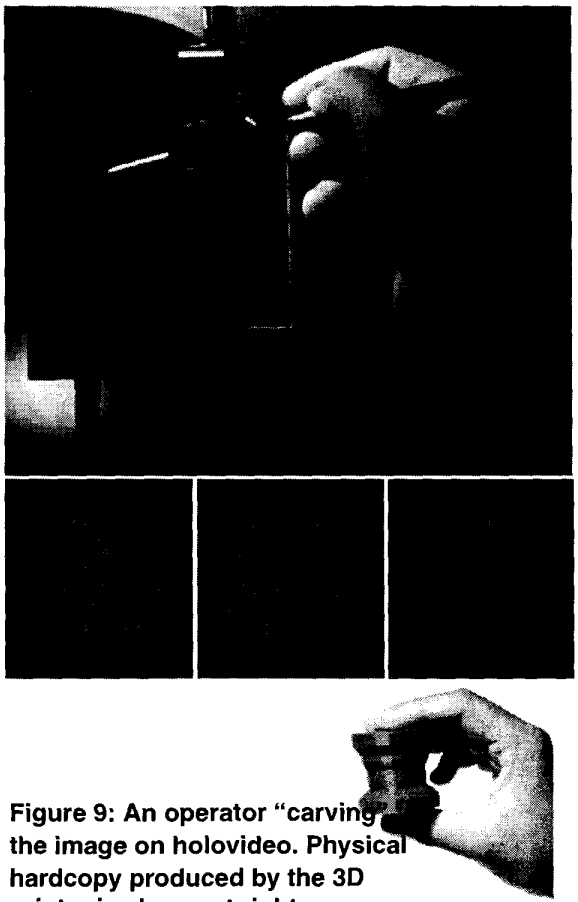


Figure 9: An operator "carving" the image on holovideo. Physical hardcopy produced by the 3D printer is shown at right.

Since the visual display is holographic, the full range of horizontal parallax is always available in the viewzone; no lag is encountered with motion of the operator's head. Additionally, no special eyewear is necessary to perceive the stereo information.

4.5.2 Differences in Visual and Haptic Renderings. Our haptic simulation models a spinning surface of revolution, but the visual representation does not spin. In order to represent a spinning holographic image, we need to be able to update all the hololines spanned by the image at a reasonable rate. As mentioned above, our system currently suffers a low frame rate with the update of only 32 lines; thus we chose to forgo animating the spinning of the holographic surface. When visual update can be more rapid, this visual animation should be included.

When the stylus tip is touched to a detail on the holographic image, touch, stereopsis and horizontal motion parallax reinforce the perception that the stylus and the holographic surface detail are spatially co-located. However, as is the case for all HPO holograms, the lack of vertical parallax causes a slight vertical shift that increases with image depth to accompany vertical head motion.

4.5.3 Differences Between Simulation and Real Task. Differences between the haptic feedback in our simulation and the feeling of carving on an actual lathe are important to note. Among them are that the simple material properties we currently simulate are quite different from those of wood or metal moving against a cutting tool. Additionally, since a "cut" applied at an instantaneous position on the cylinder surface results in a surface modification that extends around the *entire* shape circumference, the user does not experience the feeling of continuously removing material as the shape spins under the stylus. Of course, one obvious departure from reality is the 90° change in orientation of the lathe axis.

4.5.4 Sensory affordances and conflicts. The sensory conflicts in this display include all those found in the static holo-haptic displays. Additionally, since holovideo is an HPO (and therefore astigmatic) display, accommodation and convergence are disjoined to some extent (the authors are not currently aware of any systematic study to determine the behavior of accommodation when a person is viewing HPO holographic stereograms).

At the moment when an operator feels that the stylus tip is in contact with the surface, if the tip is seen either penetrating the surface or not making contact at all due to misregistration of the visual and haptic output, the visual discrepancy is striking. Due to the lag present in the holovideo pipeline, our simulation is vulnerable to this problem when the operator is actively carving the surface. Like static haptic holograms, the display does afford binocular disparity, motion parallax and pictorial cues to depth and layout. Unlike stereoscopic, half-silvered mirror displays, the hand and haptic apparatus are able to occlude the holographic image. Thus, unless hologram reconstruction is blocked, occlusion relationships are correct in the visual scene.

5. Discussion

The dynamic and the static systems described in this paper offer interaction with a holographic images on the table-top; this marks a long-held goal in the field of holography. In both of these systems, holographic images in the manipulatory space are accompanied by real objects as well (at very least the hand and haptic apparatus). In the resulting mixed-reality setting, visual, haptic and physical behavior differences between the holographic image and juxtaposed physical objects can be quite striking.

Even if we have done our best to render the holographic images with a solid, three-dimensional appearance, discrepancies between spatial images and real objects call attention to the boundary between simulation and reality. Noticeable distinction between real and synthetic objects may not necessarily impact performance in this space, but to the extent that we want to render a *physically believable* scene, we need to consider the underlying issues more carefully.

Based on observations in our laboratory and discussions with users of our systems, we have compiled a preliminary set of guidelines for generating physically believable visual-haptic displays in mixed-reality settings. We suggest that physical believability depends on how well the stimuli representing a simulated object would correspond to stimuli generated by an actual physical instantiation of that object. Rendering methods and display characteristics are obviously important factors. Additionally, all sensory modalities employed in a spatial display should act in concert to model some basic rules that, based on our experience, physical objects usually obey. To begin qualifying this topic, we group these guidelines into *display*, *rendering*, and *modeling* factors, for presenting physically believable multimodal simulations in coincident workspaces:

Display factors

Simulated and real objects should appear with the same luminance, contrast, spatial resolution, color balance, and clarity.

Visual and force images of *objects* should have "stable" spatial and temporal properties (no perceptible temporal intermittence, spatial drift, or wavering).

No time lag should be detectable between a user's action and the multi-modal response or effect of that action in the workspace.

A viewers awareness of display technology should be minimized.

Rendering factors

Computer graphic rendering or optical capture geometry should match the system viewing geometry.

Illumination used in simulated scenes should match that in the real scene (simulated shadows and specular reflections should not behave differently.

Optical and haptic material properties, as represented, should be compatible (a surface that *looks* rough shouldn't *feel* soft and spongy).

Modeling factors

The volumes of simulated objects should not interpenetrate with real or other simulated objects.

Occlusion, stereopsis, and motion parallax cues should report the same depth relationships.

Convergence and accommodation should provide compatible reports of absolute depth.

Accommodation should be permitted to operate freely throughout the volume of a simulated scene.

The range of fusion and diplopia should be the same for simulated and real scene.

All multisensory stimuli should appear to arise from a single source, and should be in precise spatial register.

Undoubtedly, more issues remain to be added to this list; the factors noted above already prescribe high technological hurdles for visual and haptic display designers.

6. Ongoing and future work

To improve the "feel" of our holo-haptic systems, we are developing new materials simulation and modeling a more realistic haptic representation of carving. To improve rendering quality on holovideo, we are developing algorithms for computing smooth-shaded and visually textured holographic images. Also, to improve our visual display update rate, we are modifying our pipeline to write hologram lines directly to the memory module.

Obviously, the performance of our dynamic holo-haptic system is limited by the available computation and communication bandwidth. This issue raises its own set of problems, some of which will no doubt be eliminated as bandwidth becomes abundant. There are other issues, such as computed holographic image quality, which merit further investigation and will continue to occupy the minds of researchers in the field of holographic imaging.

7. Conclusion

We have described two interactive multimodal spatial imaging systems in this paper. The first was a static system which allowed an operator to inspect a free-standing static holographic image with a force feedback stylus. In the second system, the operator was allowed to modify the holographic image by carving. The component haptic and holographic subsystems were described and the implementation of the whole system was detailed in each case. The sensory conflicts encountered in a coincident display format were discussed. A preliminary set of guidelines was suggested for presenting physically believable visual-haptic simulations in coincident workspaces.

8. Acknowledgments

We would like to thank the Honda R&D Company, NEC, IBM, the Digital Life Consortium at the MIT Media Lab, the Office of Naval Research (Grant N0014-96-11200), Mitsubishi Electric Research Laboratories, and the Interval Research Corporation for their support of this research. We would also like to acknowledge valuable conversations with members of the Spatial Imaging Group and the community at the MIT Media Laboratory, especially Professor Stephen Benton, Professor Hiroshi Ishii, Carlton Sparrell, Michael Halle, and John Underkoffler. We thank Adam Kropp and Benjie Chen for extending and improving the holographic design, preview and printing software. We also thank Michael Klug for his assistance in the holography lab and Yael Maguire for his 3D printing effort.

9. References

[1] Wiegand, T.E.v., The Virtual Workbench & the Electronics Training Task. Internal memo, 1994. VETT

project at MIT/RLE, 77 Massachusetts Ave. MIT, Cambridge MA 02139-4307, http://mimsy.mit.edu/.

[2] Yokokohji, Y., Hollis, R.L., Kanade, T., (1996). Vision-based Visual/Haptic Registration for WYSIWYF Display. International *Conference on Intelligent Robots and Systems,* pp. 1386-1393.

[3] Deering, M., (1992). High Resolution Virtual Reality. Proceedings SIGGRAPH'92, *Computer Graphics*, Vol. 26, No.2, pp. 195-202. July.

[4] Ishii, H. and Ullmer, B., (1997). "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms," *(CHI '97), ACM,* Atlanta, pp. 234-241, March.

[5] Wellner, P., Mackay, W., and Gold, R., (1993). "Computer Augmented Environments: Back to the Real World," *CACM*, Vol. 36, No. 7, July.

[6] Servos, P., Goodale, M.A., Jakobson, L.S., (1992). The role of binocular vision in prehension: a kinematic analysis. *Vision Research.* Vol 32, No. 8, pp. 1513-1521.

[7] Marotta, J.J., Goodale, M.A. (1997). Elevation in the visual scene: calibrating a monocularly guided reach. *Invest. Opthal Vis Sci* 38:988.

[8] Marotta, J.J., Kruyer, A., Goodale, M.A., (1998). The role of head movements in the control of manual prehension. *Exp Brain Res* 120:134-138.

[9] Ellis, S.R., Bucher, U.J., (1994). Distance Perception of Stereoscopically Presented Virtual Objects Optically Superimposed on Physical Objects by a Head Mounted See-Through Display. *Proceedings of the 38th Annual meeting of the Human Factors and Ergonomics Society.* Nashville, TN.

[10] Traub, A.C., (1967). Stereoscopic Display Using Varifocal Mirror Oscillations. *Applied Optics*, 6,6, pp. 1085-1087, June.

[11] Halle, Michael, (1997). "Autostereoscopic Displays and Computer Graphics" *Computer Graphics* (A publication of ACM SIGGRAPH) Volume 31, Number 2, May.

[12] Halle, M.H., (1994)."Holographic stereograms as discrete imaging systems," in: S.A. Benton, ed., *SPIE Proc.* Vol. #2176: *Practical Holography VIII,* pp. 73-84.

[13] Halle, M., Kropp, A., (1997). "Fast Computer Graphics Rendering for Full-Parallax Spatial Displays". in: S.A. Benton, ed., *SPIE Proc. Practical Holography XI.*

[14] Klug, M., *et al.* (1997), Optics for full-parallax holographic stereograms, in S.A. Benton, ed., *Proceedings of the IS&T/SPIE's Symposium on Electronic Imaging, Practical Holography XI.*

[15] W.J. Farmer, *et al.*, (1991). Application of the edge-lit format to holographic stereograms, in: S.A.Benton, ed., SPIE Vol. 1461, Practical Holography V pp. 215-226

[16] St. Hilaire, P., (1994). "Scalable Optical Architectures for Electronic Holography", Ph.D. Thesis, MIT Program in Media Arts and Sciences, Massachusetts Institute of Technology.

[17] Jones, M.R.E., (1994). The Haptic Hologram, *Proceedings of SPIE, Fifth International Symposium on Display Holography,* Vol. 2333, pp. 444-447.

[18] Plesniak, W., Klug, M., (1997). Tangible holography: adding synthetic touch to 3D display, in S.A. Benton, ed., *Proceedings of the IS&T/SPIE's Symposium on Electronic Imaging, Practical Holography XI.*

[19] Pappu, R., *et al.*, (1997). A generalized pipeline for preview and rendering of synthetic holograms, in S.A. Benton, ed., *Proceedings of the IS&T/SPIE's Symposium on Electronic Imaging, Practical Holography XI.*

[20] Underkoffler, J., (1991). "Toward Accurate Computation of Optically Reconstructed Holograms", S.M. Thesis, Media Arts and Sciences Section, Massachusetts Institute of Technology.

[21] Watlington, J., *et al.*, (1995). A hardware architecture for rapid generation of electro-holographic fringe patterns, in S.A. Benton, ed., *Proceedings of the IS&T/SPIE's Symposium on Electronic Imaging, Practical Holography IX.*

[22] Mark, W.R., *et al.*, (1996). Adding Force Feedback to Graphics Systems: Issues and Solutions. *Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH,* pp.447-452.