

# Coincident Display Using Haptics and Holographic Video

Wendy Plesniak and Ravikanth Pappu  
Spatial Imaging Group, MIT Media Laboratory  
Cambridge, MA  
{wjp,pappu}@media.mit.edu

## ABSTRACT

In this paper, we describe the implementation of a novel system which enables a user to “carve” a simple free-standing electronic *holographic* image using a force-feedback device. The force-feedback (or *haptic*) device has a stylus which is held by the hand like an ordinary cutting tool. The 3D position of the stylus tip is reported by the device, and appropriate forces can be displayed to the hand as it interacts with 3D objects in the haptic workspace. The haptic workspace is spatially overlapped and registered with the holographic video display volume. Within the resulting *coincident visuo-haptic workspace*, a 3D synthetic cylinder is presented, spinning about its long axis, which a person can see, feel, and *lathe* with the stylus. This paper introduces the concept of coincident visuo-haptic display and describes the implementation of the lathe simulation. After situating the work in a research context, we present the details of system design and implementation, including the haptic and holographic modeling. Finally, we discuss the performance of this prototype system and future work.

## KEYWORDS

Haptics, holography, electro-holography, autostereoscopic display, offset display, coincident display.

## INTRODUCTION

To recognize the intimate dialog between materials and the skilled eyes, hands, and intuition of the craftsperson is to acknowledge the enormity of the technology and interaction design tasks which still lie ahead of us. Ideally, we would rally the full exploratory and manipulative dexterity of the hand, and the rich sensory capabilities of both hand *and* eye to the tasks we engineer for.

Consider the domain of *traditional craft*, in which gaze and touch convene in the same location: vision directs the hand and tool; the hand senses, manipulates tools and coaxes material to take an envisioned form. Such tight alliance of eye and hand has traditionally been fundamental to tasks in which material is artfully worked into form, and a similar condition may hold for other domains as well, like surgery, component assembly, or repair and maintenance training.

Yet, in most computer-assisted applications, the hands manipulate a pointing device while the gaze is turned to a screen. Such *offset* display configurations, which direct eyes

displayward and hands to controllers elsewhere, comfortably and naturally facilitate some activities (like driving a car, playing a musical score, or moving a cursor with a mouse). In such familiar manual tasks, vision is useful for transporting the arm/hand to an object, but manipulation can often proceed quite well either in the absence of vision (after physical contact is made) or with the monitoring of visual feedback provided elsewhere. However, this may not be the best paradigm for all tasks—especially, those which are harder to control, require constant and precise visual and haptic monitoring and near-constant manual response.

In this paper, we describe an early prototype system which *spatially reunites* the focus of eye and hand and also takes a step toward bringing materials-working pleasure to computer-assisted design. While there are several conventional kinds of visual display hardware suitable for coincident visuo-haptic display of 3D information—head tracked LCD shutter glasses or head mounted displays (HMDs) combined with stereo computer graphics for instance—and while many of these visual display options currently offer adequate image quality and frame rate, they are cumbersome to wear and have attendant viewing problems. Instead, we are using a prototype glasses-free *autostereoscopic display* which allows untethered movement throughout the view-zone.

This prototype display device, MIT’s second-generation holographic video (*holovideo*) system, is capable of rendering moving, monochromatic, free-standing, three-dimensional holographic images. Currently, this device has its own shortcomings, but many will be addressed by future research and routine advances in technology. For position tracking and force display, 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. The haptic and visual workspaces are physically co-located so that a single, free-standing multimodal image of a cylinder to be “carved” is presented.

In the coincident workspace, a user can see the stylus interacting with the holographic image while feeling forces that result from contact with its force model (Figure 1). As the user pushes the tool into the simulated cylinder, it deforms in a non-volume-conserving way and an arbitrary surface of revolution can be fashioned. Ultimately, the finished computer model can be dispatched to a 3D printer providing an actual hardcopy of the design (Figure 5b). In effect, a user “sculpts light” and produces a physical result.

With these combined apparati and supporting computation, we are beginning to investigate high-quality multimodal dis-

play and interaction that is more *Newtonian* than *symbolic*, which may be preferable for tasks which have traditionally been practiced in this fashion.

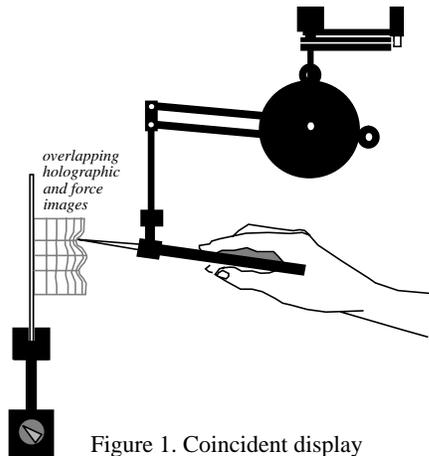


Figure 1. Coincident display

## BACKGROUND

One well-established approach to joining the eyes and hands in a coincident workspace is to employ manipulable "wired" physical objects as controllers for digital objects or processes. Several research efforts are investigating the use of physical handles to virtual objects by attaching interfacing sensors or other electronics to real objects. These *tangible* objects then act as *physical controllers* for virtual processes, providing whole-hand interaction and rich visuo-haptic feedback that seems both natural and obvious. In these applications, a participant perceives his or her own body interacting with physical interface objects, but usually *also* monitors the action-outcome on another separate display or in the ambient environment.

One such project, called Graspable User Interface: Bricks [12], employed basic physical objects called "bricks" which were physical instantiations of virtual objects or functions. Once a brick was attached to a virtual object, the computational model became itself *functionally graspable*. A brick might be used, for instance, to geometrically transform a virtual object to which it was attached, availing direct control through physical handles. Tactile and kinesthetic feedback are also present and exploitable with such an interface; thus the ability to operate quickly and efficiently, using two-handed input is possible. Extending this work to incorporate a small set of differentiable geometries and material textures among the bricks could increase a person's ability to identify and manipulate the bricks without having to look at them. This could afford visual attention the luxury of continuously monitoring the application state displayed elsewhere.

Two additional systems which give primacy to both eyes and hands in manipulatory space should be mentioned; one is Wellner's Digital Desk project at EuroPARC [14] and the other is Ishii's metaDESK project at the MIT Media Laboratory [13]. Both provide desktops on which physical and digital information commingle, and accept either gestures made with hands / pencils (Digital Desk) or the manipula-

tion of active objects (metaDESK) as input.

The Digital Desk project represents an attempt to render the computer desktop onto a real desk surface, and to merge common physical desk-objects with computational desktop functionality. The system employs a video projector situated above the desk for display of information, and a nearly co-located camera to monitor a person's movements in the workspace. Hand gestures are interpreted by a computational vision algorithm to be requests for various utilities that the system offers.

The metaDESK project attempts to physically instantiate many of the familiar GUI mechanisms (menus, windows, icons, widgets, *etc.*) in the form of tangible user interfaces (TUI's). The mapping between physical icons and virtual ones can be literally or poetically assigned; for instance placing a small physical model of MIT's Great Dome on the desk surface might cause an elaborate map of MIT to be displayed. In addition to summoning the map to the display and indicating its position, the physical Great Dome icon can be moved or rotated to correspondingly transform the map.

The metaDESK system design includes a flat rear-projected desk surface, physical icons and functional instruments for use on the surface. The state of these physical objects is sensed and used as application input. Not only can the state of virtual objects be changed by manual interaction with physical objects, but part of the display itself can be "hand-held" and likewise manipulated. The metaDESK project underscores the seemingly inexhaustible palette of ideas for instrumenting interactive space, harkening to the rich set of sensibilities and skills people develop from years of experience with real world objects, tools, and their physics.

A wide variety of virtual reality (VR) and augmented reality (AR) application areas such as telesurgery, maintenance repair and training, computer modeling and entertainment employ haptic interaction and high-quality computer graphics to study, interact with or modify data. Here, many applications employ *instrumented* force-feedback, rather than physical objects and whole-hand interaction, and trade off sensory richness for flexibility in physical modeling and visual / force rendering.

Most existing applications offset the visual and manual workspaces, but several diverse efforts to conjoin eye and hand in interactive applications exist. An example thematically related to our work is the compelling "Virtual Lathe" described and presented at the SIGGRAPH'92 conference by Michael Deering [10]. In this demonstration, a head-tracked stereo display showed a virtual stock, spinning about its long axis, which a person could interactively lathe using a 3D mouse in the shape of a rod. The demonstration required a participant to wear LCD shutter goggles for stereo viewing, but had no provision for force feedback.

A very interesting system which incorporates force feedback but no stereo viewing has been demonstrated by researchers at Carnegie Mellon University [9] called the

WYSIWYF (What You See Is What You Feel) display. The visual display behaves like a moveable “magic window”, interposed between the viewer’s eyes and hand, and through which the hand can be seen interacting with a virtual, tangible scene. The work employs a six degree-of-freedom haptic manipulator and monographic visual rendering to combine three pieces of information in this final coincident display: a video image of the operator’s hand/arm, the computer graphically rendered scene, and the accompanying force model. The visual display is a color LCD panel with a CCD camera attached to its backplane. This display/camera unit can be moved with respect to the physical scene, while vision-based pose estimation is employed to determine its new orientation. The display shows a computer graphic view of the synthetic scene generated from the newly-computed viewpoint, composited with a live Chroma Keyed image of the operator’s hand/arm moving behind the display and interacting with the haptic device. This display cannot currently reproduce correct occlusion relationships between the hand/arm and virtual objects and provides only monocular cues to scene depth (no stereo viewing or head-tracked motion parallax is available).

In other systems which employ a coincident workspace, the use of a half-silvered mirror to combine an image of the CRT’s pixel plane and the haptic workspace is a historically popular and frequently used technique. One such example is the “Virtual Workbench” [8], developed at the The Virtual Environment Technologies for Training (VETT) Group at MIT’s Research Lab for Electronics. This system, used to study human sensorimotor capabilities and to develop training applications, employs a Phantom haptic interface and the half-silvered mirror technique with stereo computer graphics for coincident 3D display. Representing correct occlusion relationships between the hand and simulated objects is a problem in this display configuration too. Additionally, the workspace that can actually be shared by the visual display and the hand is depth-limited in stereoscopic systems; inherent in these displays is an accommodation-convergence mismatch—a functional disengagement of several systems of the eye which normally function in cooperation. If scene depth is not designed well for the display’s particular viewing geometry, eye strain, headaches and unfuseable stereo images can result. Of course, the very purpose of combining the manual and visual workspace is to visually monitor the hand (or hand-held tool) and its interaction with the object or material. Consequently, the ability to keep both the displayed object and the hand in visual focus is essential, and careful design must be employed to render it so.

Holographic displays eliminate this particular design problem by permitting a viewer to freely converge and accommodate to any point in the display volume. The combination of haptics and holography was first investigated by researchers at De Montfort University for an object inspection task [1]. 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. Similar informal experiments in combining reflection transfer holo-

grams with force-feedback were also performed by researchers at the MIT Media Laboratory’s Spatial Imaging Group. In all of these efforts the interacting hand could literally block the reflection hologram’s illumination and prevent image reconstruction.

This problem was addressed by employing full-parallax edge-illuminated holograms in combination with the Phantom for the inspection of static 3D models [4]. The edge-illuminated hologram format allowed hand movements in any part of the visual workspace. Thus a viewer could haptically explore the spatially-registered force model while visually inspecting the holographic image details over a wide field of view. All of these displays were static, however; no dynamic modification could be made to the data presented.

### SYSTEM DESIGN

Marrying haptics and holovideo permits us to render simple *dynamic* scenes in the user’s manipulatory space, the domain of real objects. 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 spinning cylinder by the user’s hand, and requests hologram updates to local regions of the visual display where changes have occurred. The haptics and hologram modules rely upon separate and characteristically different representations of the cylinder, which are carefully spatially and metrically registered. From the point of view of the 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 overview is shown below in Figure 2.

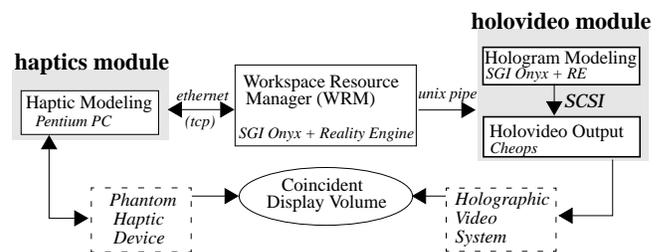


Figure 2. System overview

### Haptic Modeling

Research in haptics modeling is yielding methods to model the surface and bulk properties of materials, and the computational forces generated as we mechanically interact with them. The fidelity with which computational haptics is currently able to render both the pleasing *feel* of a material interacting with our tools, and the mechanical *cues* that relay information about object and material integrity is rapidly progressing.

Our haptic device can display force to the user's hand according to its position-based interaction with the computational models describing the object's geometry, bulk and tactual properties. Six encoders on the device provide positional information resolved to approximately 0.1 mm, and three servo motors provide force display up to roughly eight Newtons, within a workspace of about 290 x 400 x 560 mm<sup>3</sup>.

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 1 gram, 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 run on a Pentium PC with an average servo rate of 1500Hz.

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  $\Delta$ ) the surface deforms. A very simple method for surface deformation is used: the two control points straddling the penetration location are displaced toward the central cylinder axis by a fraction of the penetration distance. The upper point is displaced by  $tk\Delta$ , and the lower by  $(1-t)k\Delta$ , 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  $k\Delta$ . Thus, control point displacement modifies the circumference of the cylinder at height  $h$ , as force is interactively applied.

The parameters  $k$  and  $\Delta$  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.

## Holographic Video Modeling

We employ the second generation of holovideo in this work [5, 6]. This system is capable of displaying monochromatic, horizontal-parallax-only (HPO) images in a volume of 150 x 57.5 x 150 mm<sup>3</sup>, 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. 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 [7].

## IMPLEMENTATION

### Haptics Module

The Workspace Resource Manager (WRM) running on the SGI/Onyx 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 server 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.

Any 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 0.033 second. 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 6.9 mm high within an update interval. Though this assumption usually puts us within the realm of normal interaction speed, eventually, communicating a variable

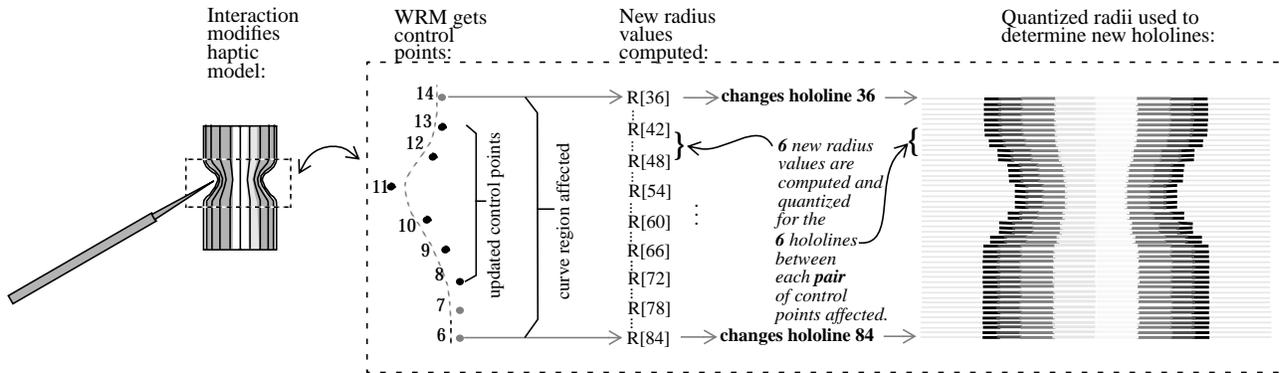


Figure 3. Determining which display lines are affected by changes in haptic model

number of control points to reflect the precise region of change would be more robust, and future work will implement this change.

### 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 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,  $\mathbf{R}$ , 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 3). These new radius values are reflected in the state vector  $\mathbf{R}$ . 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 [3], 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 47.9 mm high. These holograms, from largest to smallest radius, are loaded sequentially into the Cheops memory module. It would be possible to compute a fewer *total* number of lines if we omitted visual texture from the object or restricted texture to be periodic. 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  $\mathbf{R}$  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.

### 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 4). The final holographic image is assembled using hololines 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.

## RESULTS AND DISCUSSION

### Simulation Results

When an operator carves the holographic surface of revolution with the Phantom, the hologram image changes

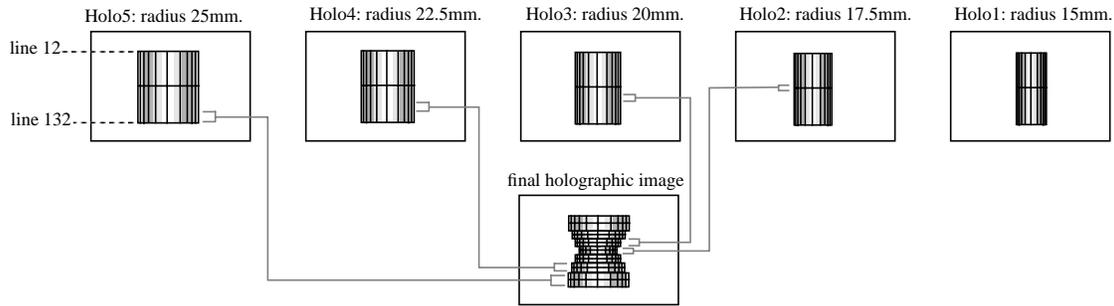


Figure 4. Method of assembling final holographic image

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 5.

### 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 [11]. 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.

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.

### Differences in the 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 choose to forgo animating the spinning of the holographic surface. When visual update can be more rapid, this visual animation will eventually be included.

When the stylus tip is touched to an image detail on the holographic surface, touch, stereopsis and 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.

### Haptic 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.

### Sensory Conflicts

Unless idiosyncratic physics are intentionally being modeled, all sensory modalities employed in a spatial display should act in concert to depict some basic rules that, based on our experience, physical objects usually obey. We have observed some important rules, and undoubtedly more still need to be considered:

- ◆ the volumes of real or virtual objects should not interpenetrate,
- ◆ occlusion, stereopsis, and motion parallax cues should report the same depth relationships,
- ◆ visual and force images of *objects* should have "stable" spatial and temporal properties (no perceptible temporal intermittence, spatial drift, or wavering)
- ◆ optical and haptic material properties, as represented, should be compatible (a surface that *looks* rough shouldn't *feel* soft and spongy),
- ◆ all multisensory stimuli should appear to arise from a single source, and all should be in precise spatial register.

If *intermodality* or *intramodality* sensory cues yield conflicting information about the physical location and

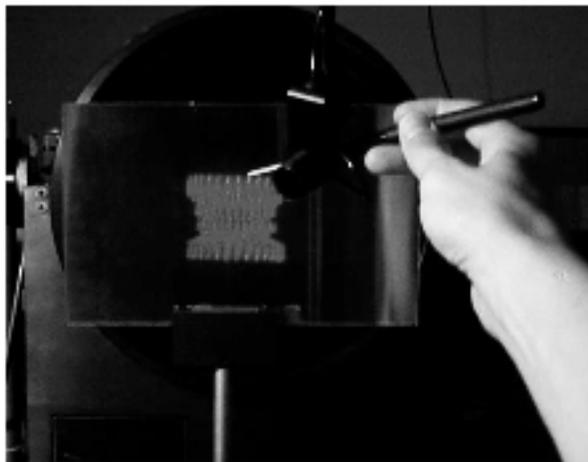
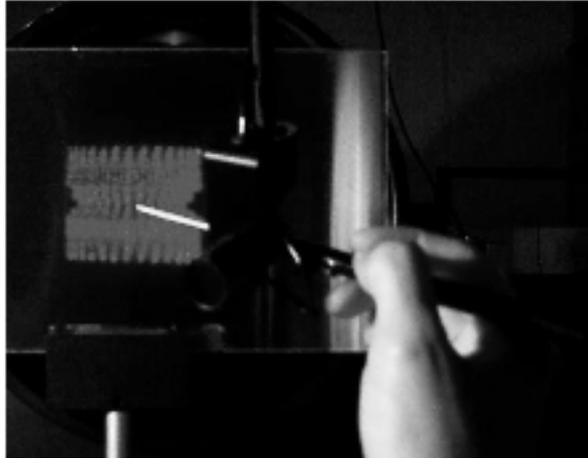


Figure 5a. Interactive carving in the coincident workspace



Figure 5b. Physical hardcopy from a Stratasys 3D printer

boundaries of an object, about the nature of its material properties, or about its relative distance from the observer, the impression of the simulation as a single multimodal event can be compromised. Two conflicts which are most troublesome are *spatial misregistrations* and *occlusion violations*. 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.

Allowing interaction between the output of optical projection systems, like holograms, and an operator's hands (plus physical objects or instruments), permits object depth relationships to occur which violate occlusion rules obeyed by the physical world. Normally, when we see the image of one object blocking the image of another from view, we understand that the occluded object is farther from our eyes. In our system, it is possible to interpose part of the stylus between the holographic image and physical output plane of the holovideo optical system, thus blocking the image from the line of sight between the viewer and stylus. In this event, it appears that the farther object (stylus) occludes the nearer (holographic image). This anomalous cue is strong enough to confuse perception, even when correct depth reporting from stereopsis and motion parallax is available.

#### ONGOING / FUTURE WORK

Currently, we are working on improving the fidelity of materials simulation, and modeling a more realistic haptic representation of carving. We are developing algorithms for computing smooth-shaded and visually textured holographic images. Also, we are modifying our pipeline to write hologram lines directly to the memory module to increase our visual display update rate. When we can update the holovideo display more rapidly, our general results will improve markedly and simulating more complicated dynamics becomes a tractable pursuit.

It is important to note that evaluating affordances in a system that spatially unites manual interaction and visual display, and determining whether performance (for some tasks) is *truly enhanced* in such a workspace requires a rigorous comparison between offset and coincident display formats. While it is premature to consider performance testing or user preference evaluation using the system described here, a controlled study could be made presently by using, for instance, a more conventional stereoscopic display and a Phantom.

Certainly, electronic holography and force feedback technologies still present us with many fundamental engineering challenges, and evaluation of their mutual efficacy should be carried out after these challenges have been addressed.

#### CONCLUSION

We have described an interactive multimodal spatial imaging system in this paper. We have situated the system in a research context and outlined the requirements of such a

coincident workspace system. The component haptic and holographic subsystems were described and the implementation of the whole system was detailed. Both the benefits and problems raised by using a coincident display format were discussed.

Our belief in the importance of high-quality visual and haptic cues as well as the attentive dialog between vision, hands, tools, and the material they manipulate is at the root of the work presented here. The coincident-workspace system takes a step toward giving primacy to the skilled eyes and hands of the craftsman and offers, for further inquiry, a possible interaction technology for future digital design studios.

#### ACKNOWLEDGEMENTS

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 also thank Bill Verplank for inspiration and valuable discussion, and Yael Maguire for his 3D printing effort.

#### REFERENCES

1. Jones, M.R.E., The Haptic Hologram, *Proceedings of SPIE, Fifth International Symposium on Display Holography*, Vol. 2333, pp. 444-447, 1994.
2. McCullough, M., Abstracting Craft: the practiced digital hand, the MIT Press, Cambridge, MA, 1996.
3. Pappu, R., *et al.*, 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*, 1997.
4. Plesniak, W., Klug, M., 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*, 1997.

5. St. Hilaire, P., "Scalable Optical Architectures for Electronic Holography", Ph.D. Thesis, MIT Program in Media Arts and Sciences, Massachusetts Institute of Technology, 1994.
6. Underkoffler, J., "Toward Accurate Computation of Optically Reconstructed Holograms", S.M. Thesis, Media Arts and Sciences Section, Massachusetts Institute of Technology, 1991.
7. Watlington, J., *et al.*, 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*, 1995.
8. 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/>.
9. Yokokohji, Y., Hollis, R.L., Kanade, T., Vision-based Visual/Haptic Registration for WYSIWYF Display. *International Conference on Intelligent Robots and Systems*, pp. 1386-1393, 1996.
10. Deering, M., High Resolution Virtual Reality. *Proceedings SIGGRAPH'92, Computer Graphics*, Vol. 26, No.2, pp. 195-202. July 1992.
11. Mark, W.R., *et al.*, Adding Force Feedback to Graphics Systems: Issues and Solutions. *Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH*, pp. 447-452, 1996.
12. Fitzmaurice, G., Ishii, H., Buxton, W., "Bricks: Laying the Foundations for Graspable User Interfaces," *Proceedings of Conference on Human Factors in Computing Systems (CHI '95)*, ACM, Denver, May 1995, pp. 442-449.
13. Ishii, H. and Ullmer, B., "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms," (*CHI '97*), ACM, Atlanta, March 1997, pp. 234-241.
14. Wellner, P., Mackay, W., and Gold, R., "Computer Augmented Environments: Back to the Real World," *CACM*, Vol. 36, No. 7, July 1993.