Emancipated Pixels: Real-World Graphics In The Luminous Room

John Underkoffler, Brygg Ullmer, and Hiroshi Ishii
Tangible Media Group, MIT Media Laboratory

ABSTRACT

We describe a conceptual infrastructure – the *Luminous Room* – for providing graphical display and interaction at each of an interior architectural space's various surfaces, arguing that pervasive environmental output and input is one natural heir to today's rather more limited notion of spatially-confined, output-only display (the CRT). We discuss the requirements of such real-world graphics, including computational & networking demands; schemes for spatially omnipresent capture and display; and issues of design and interaction that emerge under these new circumstances. These discussions are both illustrated and motivated by five particular applications that have been built for a real, experimental Luminous Room space, and by details of the current technical approach to its construction (involving a two-way optical transducer called an *I/O Bulb* that projects and captures pixels).

CR Categories and Subject Descriptors: H.5.2 [User Interfaces] Input devices and strategies; H.5.1 [Multimedia Information Systems] Artificial, augmented, and virtual realities; B.4.2 [Input/Output and Data Communications] Input/Output Devices – Image display; I.4.8 [Image Processing and Computer Vision] Scene Analysis – Tracking, Applications; I.3.6 [Methodologies and Techniques] – Interaction Techniques

Additional keywords: real-world graphics, luminous-tangible interfaces, projection, computer vision, architectural space, CAD

1 INTRODUCTION

For the three decades of Computer Graphics's so-far existence, the graphics in question have been confined to screens: vectorscopes, raster-scanned CRTs, head-mounted displays, and – when budget, technology, or Hollywood permits – the cinema screen. We suggest that it is time for computer-driven graphical constructs to escape the strictures of the rectangular screen and take their place as denizens of real-world architecture. Specifically, we introduce the concept of the Luminous Room: an interior architectural space whose surfaces have been made capable both of displaying visual information and of performing visual capture. By insisting on this collocated pairing of optical input and output, we make of each room surface – floor, walls, ceiling, tabletops, assorted furniture – a potential site for interaction. These are sites at which the 'room itself' can react to what's happening within its bounds, sites at which otherwise inert physical objects can become digital implements visibly surrounded by computational appendages.

This document, then, concerns the Luminous Room, dovetailing discussion of the general with the specific. Overarching questions – of appropriate technical means for achieving what we propose, of the role played by physical objects in the interactions that result, of the ways design for such a space must differ from traditional computer-visual design – are paralleled by corroborating ideas from experiments in which we have built pieces of a Luminous Room and some applications that live there.

We'll begin by presenting five such examples: working Luminous Room applications that, while hardly encompassing the full set of 'things to do with graphics in the real world', should provide background and basis for the ensuing discussion of some large-scale issues. Specifically, we explore (1) approaches to distributing visual information throughout and extracting spatial information from a room; (2) the computational and networking needs related to input, i.e. analysis of the environment; and (3) design issues attending the deployment of and interaction with computer graphics in a real-world setting.

2 EXAMPLE APPLICATIONS

One major intent of our research has been to find ways in which the facilities of a Luminous Room – pretending for a moment that the formidable challenges of its basic technical realization are already met – can be usefully employed. If every room surface really is capable of display, what interactions does it make sense to pursue there? The following five applications will serve as grounding for our later discussion of large-scale implementation issues; they also illustrate one class of problem that we've found nicely susceptible to treatment by reactive real-world computer graphics.

2.1 Illuminating Light

Designed as a rapid prototyping tool for optical engineers, *Illuminating Light* provides a collection of simple objects representing real-world optics: lenses, mirrors, lasers, beamsplitters, and recording film. The system acts to imbue these models with the same meaning and function as their real-world counterparts [13], so that placing the laser-model on the ordinary table that is the application's arena results in a graphical beam of light, apparently emanating from the laser's front aperture. This beam remains registered with the laser as it is rotated and moved about the workspace. Other optics models placed in an existing beam's path have the expected effects: a mirror-model reflects the beam; a lens-model spreads one beam into a fan of many; a beamsplitter-model reflects part of the beam and transmits the remainder; and a recording-film-model absorbs incoming beams to form an eventual image. Ancillary information – the length of each beam and the angle of beam-bounce at each mirror and beamsplitter surface – is presented as a collection of numerical annotations that float in the vicinity of the beams and optical

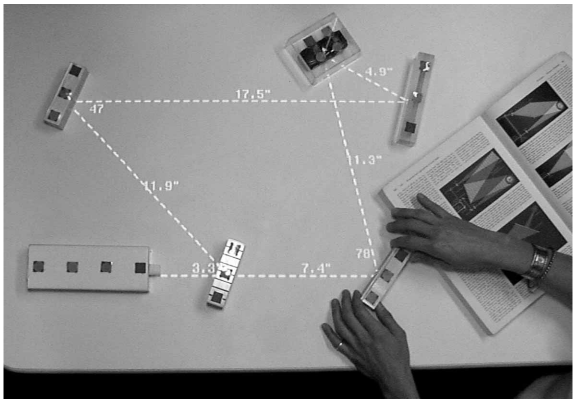


Fig 1: *Illuminating Light*: physical optics, digital light beams

---

20 Ames St. Cambridge MA 02139 / jh@media.mit.edu

SIGGRAPH 99, Los Angeles, CA USA

elements to which they refer.

Additional domain knowledge concerning holography is built into the application: as an experimenter works at arranging a hologram-recording layout, a nearby display shows optical path-length-matching information (critical for successful recording). Similarly, once a viable setup has been achieved a simulated holographic reconstruction appears in the workspace. *Illuminating Light* has aroused a great deal of interest in our local optics and physics communities. The students and professionals who have experimented with it affirm that its direct manipulation style – "like working with the real thing" – both fosters and takes advantage of the spatial understanding inherent to work with real optics. Their comments have also indicated that for many tasks the system is easier and faster to use than the on-screen CAD-style applications that are the other alternative to prototyping directly in the lab.

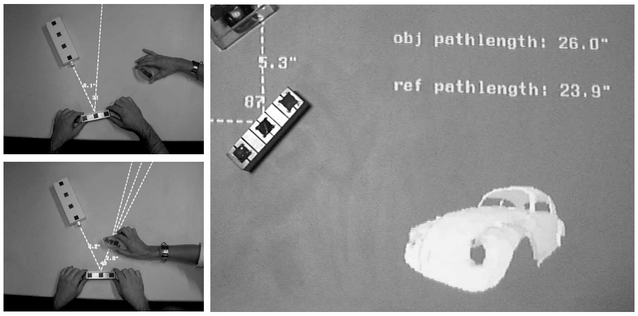


Fig 2a: collaborative optics design; 2b: simulated hologram

2.2 Chess-&-Bottle

An application that involves a vertical rather than horizontal surface, this collection of design experiments covers an entire wall in a small office. Placing a large chessboard anywhere on the wall engages a chess system that rapidly populates the board with animated pieces; moving the board induces the individual pieces to scramble into the appropriate new positions. In the same space (and at the same time, if desired), digital artifacts can be stored in a physical container. Images, numbers, text, and even regions of live video can be brought into loose association with a glass vase, which extends graphical springs to capture any of these 'documents' in its proximity. A simple gesture – rotation about its vertical axis – causes the vase to fully ingest its collection of documents. The vase can then be moved to a new location and twisted once more, whereupon (like any good physical container whose exterior location does not alter what is contained) it disgorges those same documents.

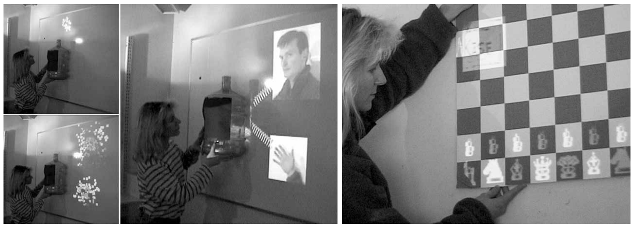


Fig 3a: physically stored digital elements; 3b: chess

Individual documents (live video windows, at the moment) are created with a colored paddle that's used as a spatial pointer. The same paddle is used to move existing documents over the wall, to bring them into or out of association with the container-vase, and ultimately to dispose of them in a physical trash can.

2.3 seep

A cellular automata system is used as the basis for *seep*, an interactive simulation that allows physical objects placed in a workspace to act as obstacles in a purely computational fluid flow. The flow is depicted as a grid of field lines whose orientation and length show the local direction and speed of fluid transport. A lattice gas laid out on a hexagonal grid – the FHP formulation [5] – expresses fluid behavior as an aggregate of individual particle motions; particle interactions are dictated by a small set of collision rules that accurately lead to the dynamics predicted by the *Navier-Stokes* equations. Physical objects placed in this graphical flow are tracked and reduced to a two-dimensional silhouette, whose interior region is then analyzed into the grid of FHP cells as a collection of obstacle cells. The result is an engaging tool for experimenting with basic fluid physics.

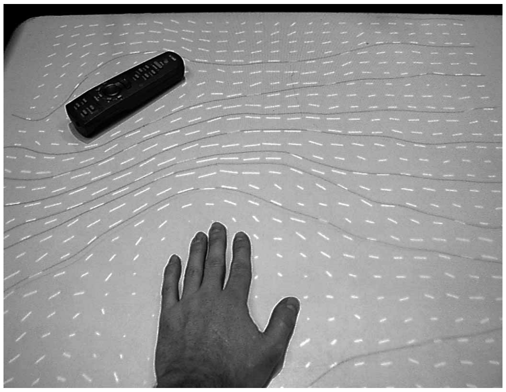


Fig 4: *seep*'s fluid diverted by physical obstacles

2.4 Urp

We have built a second 'professional' application, expanding on many of the ideas we'd been exploring with *Illuminating Light*, to address the field of urban design and planning. The resulting system, called *Urp*, also permits direct manipulation of basic objects – in this case architectural models – to affect an underlying simulation. One part of this simulation attaches computational solar shadows to each building model in its purview. The buildings are continuously tracked and each shadow is closely registered to the structure that is its source, so that a convincing illusion of the shadows' authenticity results for many experimenters (despite the incongruity of a wireframe construction casting a solid shadow!).

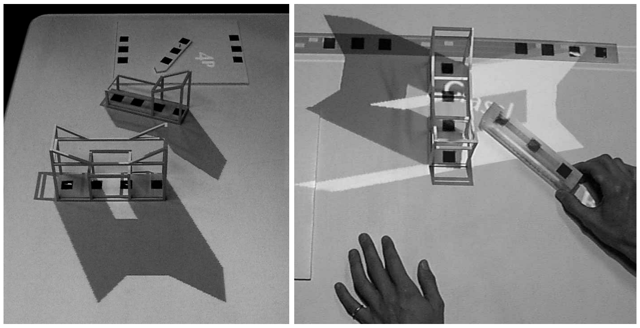


Fig 5a: shadows; 5b: material wand makes a building glass

But where *Illuminating Light* provides essentially a single class of object – i.e. each object is a stand-in for some optical element, designed to closely emulate the behavior of that real-world counterpart – *Urp* expands the repertoire. So while the architectural models are in their 'literalness' clearly analogous to the earlier optics models, *Urp* also provides tool-objects: objects that act on other objects in the simulation, or that act on the global state of the simulation. Thus, experimenters can rotate the hour hand of a simple clock-object placed on the table to change the system's time – and thereby the current solar position; the shadows swing around accordingly. This adjustable time, together with the graspable physicality of the architectural models, provides a straightforward mechanism for performing shadow studies.

Similarly, a measuring-tool can be used to 'draw' a distance measurement between any two structures; the line that represents this metric connection persists as the models are moved about the workspace, always updating the displayed distance. A material tool, brought into contact with any building, changes its 'facade' from brick to glass: the simulation then additionally displays the calculated solar reflection from each exterior surface. This makes possible planning for the interaction between glass-faced structures and nearby freeways, which can become significantly hazardous when low-angle sunlight is reflected into oncoming drivers' eyes. Finally, a wind tool placed into the workspace at some particular angle summons a windflow simulator – the same FHP lattice gas underlying *seep* – that blows from the chosen direction and takes into account the obstruction of any buildings in its path. The resulting field, as with *seep*, is projected down into alignment with the workspace, enabling straightforward wind-flow studies.

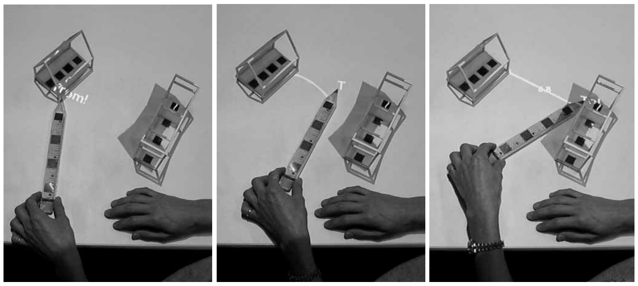


Fig 6: A measurement line is wigglingly established

Reaction to *Urp* from practicing and academic urban planners and architects has been emphatically positive. Plans are already under way to duplicate the system several times over for a new design studio in our university's architecture school. This new version will also include a distribution mechanism (discussed in §5.2.4) that allows several planners, each at a separate workspace, to simultaneously see and collaborate on an evolving design.

2.5 Distributed Illuminating Light

Finally, the ongoing expansion of our Luminous Room infrastructure has allowed the integration of multiple *Illuminating Light* workspaces. Our distribution scheme (§5.2.4, again) permits two workspaces to be physically juxtaposed, leading simply to a larger workspace. For the alternate case, in which the workspaces are not adjacent, we have implemented two distinct modes of operation. In the first, each workspace is a 'window' onto a continuous optics workbench that's rigidly isomorphic and -metric with real space. Thus, a laser shined off the edge of one table will reappear on another if aimed properly (Fig. 7). In the second mode, all workspaces are understood as instances of the same underlying space; this is the 'I-see-what-you-see' case. Here, optics models present on one table will be graphically represented on the others, but will not be directly manipulable by experimenters there. Beams, of course, are identically represented on each table; and so each experimenter views (and has a role in constructing) the same optical layout.

3 CONTEXT

3.1 Historical

The idea of incorporating live computer-generated imagery into architectural spaces is not nearly new, of course.

Fig 7: distributed *Illuminating Light*

Myron Krueger's decades-long series of experiments, including notably the many faces of *VideoPlace*, used simultaneous video projection and capture to embed the human participant (in silhouette form) in a variety of games and simulations [7]. This same video-acquired silhouette, subjected to simple geometric- and gesture-analysis, was itself also the input to VideoPlace.

MIT ArcMac offered in 1979 the notion of the 'Media Room', an office-sized space in which an entire wall was in fact a rear-projection video system [2]. The room's occupants were able to interact with wall-displayed applications like *World of Windows* and *Put-That-There* by way of a multiplicity of input mechanisms, including physical gesture (Polhemus-style magnetic tagging), verbal commands (voice recognition), visual attention (eye tracking), and a set of more ordinary buttons & joysticks.

Most inspiring and aesthetically potent is the work of Michael Naimark, who since the 1970s has worked with immersive video and film projections schemes. Among them, especially germane is his 1984 piece *Displacements*, in which a central rotating camera had earlier captured actors' antics in a mocked-up living room; the room was then painted entirely white (furniture, props, and all) and the developed film placed in a rotating projector precisely registered with the original camera. Visitors who entered the room watched a finite frame of color and life sweeps around, animating the otherwise sterile environment.

3.2 Contemporary

Viewed broadly, the Luminous Room shares certain individual aspects with a handful of other research projects, including the CAVE [3]; augmented reality systems such as Karma [4]; 'smart rooms' as in [10]; several desk- and workbench-based VR systems such as [1]; and, more recently, projective environments such as [11]. Our research differs from these systems in several key respects. We seek to 'paint on' the physical objects and surfaces that constitute the real world, bypassing the encumbering mediation of see-through displays and tethered tracking technologies. Our work is also distinguished by its reliance upon systems of physical artifacts both as representations of and as mediums of interface with the digital world.

This latter concern is shared by much of the 'Tangible Bits' work of MIT's Tangible Media Group [6], but where the 'phicons' of these systems are typically endowed with *symbolic correspondences* between digital meanings and physical manifestations, Luminous Room objects have more often demonstrated a *direct correspondence* with pre-existing physical artifacts (e.g. optics and buildings), along with a corresponding faithfulness to the original interaction modalities of these items.

Perhaps closest in spirit to our present work is Wellner's DigitalDesk system, which used projection and video-capture techniques to merge (literally) the computer desktop with the physical office desktop [14].

4 GETTING PIXELS IN AND OUT

What means are plausibly available for realizing a Luminous Room? How can display and capture actually be distributed *pervasively* throughout a room?

4.1 Display

There are three basic ways to enable display at an architectural surface. A first requires replacement or occlusion of the surface in question by a planar emissive structure, varieties of which include the hoary CRT (whose applicability seems dubious, given the significant thickness involved) and the more credible flat-panel display technologies. A second approach involves projection; rear projection is often implausible because of the large 'under-surface' distances demanded, but projection from within the space remains a possibility. Finally, just-emerging active-contrast technologies hold interesting promise. These can take the form (among others) of 'electronic toner', as with Jacobsen's E-Ink work [8]: electrically rewritable but optically passive (black or white) pixels that may be affixed to various surfaces.

4.2 Capture

Similarly, we may identify several basic schemes for the acquisition and spatial tracking of the room's contents and inhabitants. Note that, in addition to the tracking of inanimate objects and humans (or perhaps other bits of biology), we may also wish to apprehend the Luminous Room's own displayed graphics – e.g. in order to employ auto-calibration schemes of the sort described in [11].

Approaches to capture include *tethered tagging schemes* (e.g. Polhemus) in which position and orientation information is reported for any object or human to which a 'receiver' has been attached. A second approach is 'Electric Field Sensing' (described in [12]), which is an *electronic non-contact tagless scheme* in which an object modifies an electric field generated and sensed by a fixed transceiver; the position and orientation of the interposed object emerge from calculations based on the characteristics of the received field. Another approach is video capture, followed by the application of any machine vision algorithm(s) – an *optical non-contact tagless scheme*. The final approach, an *electronic non-contact tagging scheme*, is any – like [9] – that affixes untethered tags (RF cavity tag, magnetic tag, etc.) to objects and then uses a set of sensors fixed under a surface to discover the spatial particulars of each tag.

4.3 I/O Bulb

We have established the dyad of optical input and projective output as our approach – for the moment, at least – to building Luminous Rooms. This decision is multiply predicated. Critically for actual implementation, camera-in and projector-out are the schemes most clearly within the current realm of the attainable. The proposition of carpeting and wallpapering a room with flat-panel displays is not only prohibitively expensive but engenders structural (in the load-bearing sense) difficulties, leading as well to slightly smaller rooms or slightly larger houses. The present immaturity of E-Ink technology precludes its use. Polhemus-style input too is expensive and does not scale well, and its signal-wire-tethers are antithetical to the non-encumbering aims of the project. Electric field sensing techniques are not yet capable of satisfactory object disambiguation and essentially require objects to be partially conducting (flesh, usually). RF tagging approaches also suffer limits on the number and material composition of objects that can be recognized.

Equally important to our decision is a philosophical consideration: using optical input and projective output is the only configuration that is largely 'non-invasive'. Although future homes and workspaces may incorporate embedded display and sensing facilities in their surfaces as a matter of course, there is for now a strong appeal to schemes that do not require laying down extra surfaces or burrowing beneath existing ones to install electronic hardware.

The particular tack we endorse involves close physical integration of the optical input and output functions. In particular, we have proposed the evolution of the ordinary lightbulb, as follows. If we first generously consider a present-day lightbulb to be an extremely low-resolution digital projector – a 1x1 pixel(s) projector, in fact, typically also sporting a 1 bit dynamic range – then the necessary evolution is easy to see. We must increase the resolution of the bulb, so that its current angularly independent output (roughly the same intensity throughout the full 4π steradian surface that surrounds it) becomes an output of angularly dependent intensity (with, say, 8 bits of dynamic range). To this high resolution bulb structure we now also add a tiny video camera, so that the bulb not only concerns itself with the light that flows *outward* through its familiar glass shell, but also collects the light from the outside environment that flows through the shell *inward*.

The result is a two-way optical information device that we call the *I/O Bulb*. The applications presented in this paper have so far made use of various I/O Bulb mockups built from off-the-shelf components (described below in §4.4). Building a 'real' I/O Bulb is itself a tangible engineering goal, an ongoing research project that we have been pursuing with an outside industrial partner (an early prototype is shown in Fig. 13b). But the I/O Bulb is also useful as a conceptual unit. If an individual I/O Bulb is capable of treating some finite region with display and scene capture, then we can build a Luminous Room in the same way that an ordinary architectural space is illuminated: through a multiplicity of bulbs. By placing enough I/O Bulbs in a room (most in the ceiling, mimicking one traditional style of illumination) and by supporting their coordination and intercommunication, we can 'tile' the entire space with graphical interactivity – and thereby build a proper Luminous Room.

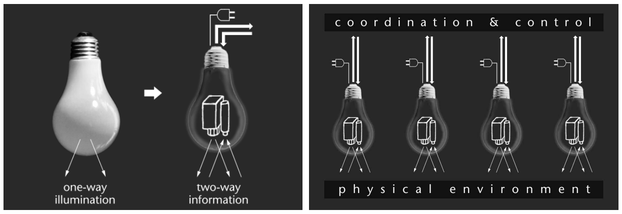


Fig 8a: from lightbulb to I/O Bulb; 8b: a Luminous Room

4.4 Projective Geometries

Several optical-geometric issues pervade any attempt to build an I/O-Bulb-like structure using off-the-shelf components.

4.4.1 Anti-Keystoning

Projector manufacturers in 1998-9 tend to assume that their products are used more or less exclusively in business presentations, and will thus only ever be placed on conference-room tables or mounted on ceilings. The anti-keystoning mechanism that is consequently integrated into every modern data projector means that its projection expands along a frustum not centered on the normal to the lens – although each parallel focus plane of this frustum is of course still properly orthogonal to the lens normal (Fig. 9). Thus a projector that points toward the horizon, straight ahead, will deposit its image well above the 'horizon-aim' of its lens, but that image will be properly rectangular and wholly in focus. This is in contrast to, say, a standard slide projector, whose projected image center will always coincide with the aim-point of its lens; moving the image higher on the wall necessitates propping up the front end of the apparatus, but the image then becomes trapezoidal ('keystoned').

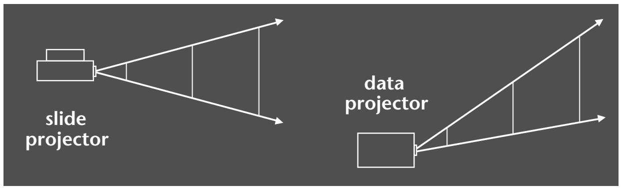


Fig 9: Normal v. 'anti-keystoned' projection

An I/O Bulb, meanwhile, is clearly only useful if the region observed by its input-camera is the same as the region painted by its output-projector. Given that available projectors cannot truly project 'forward', we are left with two prospective geometries for achieving coincidence of input and output. We can either separate the camera and the projector, so that the regions treated by each are precisely the same; or we can keep the camera and projector together (at least as close as is geometrically possible) and tip the projector downward to bring the center of the projection into alignment with the center of the camera's view. The significant shortcoming of this latter arrangement is that not only is the resulting projection trapezoidally distorted – requiring software clients to apply a final counterdistortion transformation to any imagery before display – but the plane of projective focus is now also tipped with respect to an orthogonal projection surface: correct focus is no longer possible over the extent of the image.

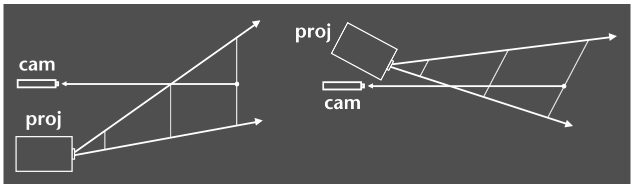


Fig 10: Geometries for aligned I/O coverage

4.4.2 Coincidence and Coaxiality

Equally serious is the issue of whether the two system components are optically coaxial or not: is there parallax between the camera's view and the projection axis? Early thought maintained that a true 'zero-parallax' (precisely coaxial, with camera and projector essentially coincident) system would be optimal; indeed there are reasons that minimizing parallax is advantageous. The fact is, however, that for an arrangement in which all objects and projections are restricted to a single plane – which frequently is the desired arrangement (e.g. an overhead I/O Bulb addressing a table surface) – the parallax issue is moot. Indeed, simply-calculated offsets can precisely account for the positions of objects that 'depart the plane', as long as the geometry of the responsible I/O Bulb is well known.

Barring the availability of a single chip with an active surface that both emits and collects light, the only way to achieve a true zero-parallax optical input-output system is through the use of a partially-silvered mirror. But a version of the I/O Bulb built early in the research using this technique quickly revealed the fundamental drawback that renders such an approach largely unworkable: optical scatter. Even an absolutely clean beamsplitter scatters a small fraction of the light that's incident upon it. Since output visibility in a normal work environment requires the projection component of the I/O Bulb to be substantially high-intensity (the more so because the beamsplitter 'throws away' part of the light that would normally reach the projection surface), a fair amount of light is unavoidably scattered from the beamsplitter surface. The camera must look through this surface as well, and whatever it might have seen of the environment beyond the I/O Bulb assembly is now drowned out by this scattered projection light.

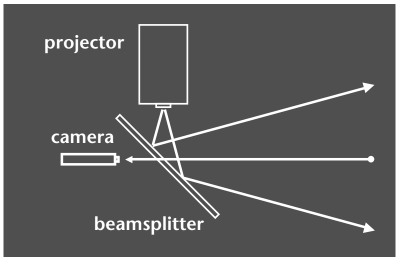


Fig 11: Scatter-crippled approach to zero-parallax

4.4.3 The Current I/O Bulb

With all this in mind we embrace (temporarily) the spatially-separated-camera-and-projector option, so that most of our functioning 'I/O Bulbs' to date have been built as shown below. Although a long-term objection to this configuration exists – we ultimately intend the I/O Bulb to be a compact device, suitable for unobtrusive and large-scale deployment – the ideological disparity is tolerable in the short term as we develop the applications that are an equally important part of the Luminous Room investigations. Too, despite our original misgivings about such a large amount of camera-projector parallax, it works quite well.

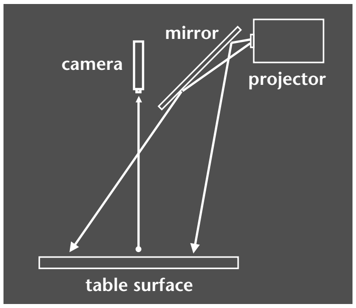


Fig 12: The currently prevalent I/O Bulb geometry

An unforeseen advantage of this physically distributed design quickly emerged. For typical 'workbench' applications in which operators stand or sit at the front of the table (i.e. on the left side in the diagram above) and at its sides, occlusion by an operator is *less* of a problem than would be the case with a zero-parallax system: shadows from an operator's hands and arms tend to be thrown forward – that is, back toward the operator herself.

5 PARSING THE ENVIRONMENT

Whether the Luminous Room acquires 'raw' information about the presence, position, and movements of the objects and people inside its confines by way of a video camera, an electric-field sensing setup, or some tagging scheme, a great challenge is always the compilation and interpretation of this low-level data into the form of 'recognized phenomena'. So, for example: a high-velocity arc has been recorded; is someone waving, or has *Illuminating Light*'s beamsplitter just been thrown in anger?

5.1 Distributed v. Centralized Approach

No matter what scheme or collection of schemes is used to implement a Luminous Room, decisions about how to distribute the resulting computational load will arise. In particular, because any currently plausible scheme involves juxtaposition of many finite-purview input and output mechanisms (i.e. no single device can 'cover the whole room'), some amount of coordination among these more basic spatial 'tiles' is required. What happens, for example, when a graphical construct needs to straddle or cross the boundary between adjacent tiles?

A central question thus concerns the granularity of the computation that reflects and is responsible for the room's tiling. The two extremes would see (1) an individual process, probably running on its own dedicated CPU, assigned to each tile, with the smallest possible amount of intercommunication keeping the tiles synchronized – viz., the fully distributed approach – or (2) a single 'omniscient' process whose massive job it is to attend to all tiles in aggregate and simultaneously; sharing of information between tiles is then not merely effortless but in fact unnecessary. In between, we may imagine a largely distributed approach that assigns individual processes to each tile but manages these via a master, supervisory process: a tack perhaps understandable as an 'Operating System for the Luminous Room'.

In §5.2.4 we present an implementation of the first (maximally fine-grained) option.

5.2 Luminous Room Vision Techniques

Here we present the fundamental choices made and techniques employed that allow our current Luminous Room implementation to track objects within its purview.

5.2.1 glimpser

Early stage, low-level vision is accomplished by the *glimpser* program, which simply identifies colored dots in its visual input. *glimpser* accepts commands from its master application to define, create, destroy, and condition 'finders'. Each finder is an independent directive to locate within the input frame a specific-sized region of some particular color. Finders, once created, can be restricted to a certain subregion of the input field, can be temporarily deactivated or fully reactivated, and can be 'de-emphasized' to be evaluated less frequently in order to streamline the search when input movements are known to be slow or very sporadic. Finally, each finder may be instructed to report only one color-spot location per frame, to report up to some fixed number of spot locations per frame, or to report fully as many spot locations as may be found per frame.

glimpser is implemented as an isolable server in which requests and directives are received and resulting reports are transmitted over a TCP/IP connection. In this way *glimpser*'s predictably heavy computational demands may be fobbed off onto another CPU altogether, leaving the 'main' CPU freely available for the full simulation and rendering needs of the application in question; or, for lighter tasks, glimpser's low-level vision efforts as well as the application-specific calculations can be assigned to the same machine. *glimpser* has been used with satisfactory results in both guises.

5.2.2 Seeing Spots

The point of this color-dot-finding is that, in most of the applications built so far for the Luminous Room, individual physical objects are tagged with unique colored-dot patterns. For a variety of reasons, not least of which is the desire to maximize reliability and stability while minimizing per-frame computational cost, we decided at the outset of all our implementation to eschew higher-level machine vision techniques (like template-matching) that attempt to identify objects through shape and other per-object attributes.

Instead, the intent is a kind of object-independent tagging scheme that – while enjoying the benefits of machine vision, like inherent multiplexing – would exhibit a special flexibility. For example, if we decide that an application needs to be able to recognize a new object, we need only declare the unique dot pattern that will be affixed to this object. Depending on the structure of the application and the intended function of the new object, this addition may not require recompilation (or indeed even restarting the application). An object-centric vision scheme would, on the other hand, typically require some form of 'retraining'. At the same time, the dot-pattern vocabulary is highly extensible, limits being imposed only by available physical space (obviously, we need the patterns to be small enough to fit on the object they identify), camera resolution, and the syntactic richness of the pattern space we establish.

An important implementation issue is the reliable isolation of genuine color dots from an unpredictable background. To wit: even with highly saturated colors chosen as pattern-dot 'primaries', the dots are at best still Lambertian reflectors. Thus there is no way to guarantee (1) that the same hue will not be present in garments, skin pigments, or unrelated objects in the environment, or (2) that brightly-illuminated surfaces in the environment won't become isomers of the dots' hues through aliasing of the CCD's chromatic response curves. So irrespective of the sophistication of *glimpser*-level algorithms, false positives will be reported and genuine dots ignored with crippling frequency. Making the dots self-luminous (say, by embedding small LEDs) could solve the problem by boosting the luminance of each to an unambiguous level in the video input field, but would uncomfortably breach the maxim that objects used by Luminous Room applications should be passive.

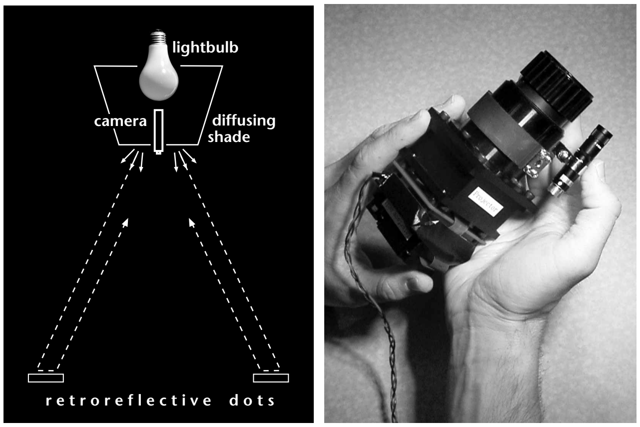


Fig 13a: retroreflective tagging; 13b: real I/O Bulb prototype

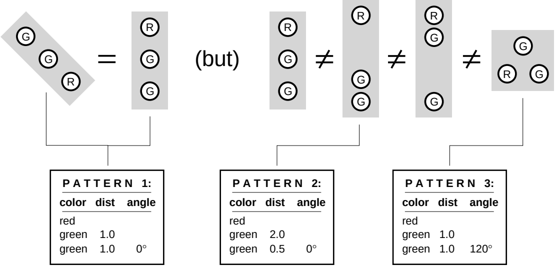
Instead, we've elected to use retroreflective dots complemented by a low-intensity, diffuse light source around the I/O Bulb's camera. A first round of dot-design employed a disk of panchromatic 3M ScotchLite material covered with a colored gel (red, green, or blue). At the same time, a moderate 60W (old-fashioned) lightbulb was incorporated into the I/O Bulb structure, placed directly above the slim video camera. A diffusive shade was constructed around the whole, with the lens of the camera protruding from the bottom.

Each dot is then illuminated by this annular diffuser and, no matter the angle of the light's incidence on it, reflects a gel-filtered version of most of this light directly back into the camera's lens. Because of this angularly selective reflection, human operators do not perceive the dots as other than normal surfaces; they seem no brighter than anything else. But from the privileged position of the camera, the dots glow fiercely: typically 2–4 stops brighter than any other part of the visual field. The critical result of all this is that it is now necessary to stop down the camera (either optically or electronically) in order to bring the high-luminance dots back within its dynamic range – and but doing so renders most of the rest of the input field black. Reliable dot isolation is thereby assured.

New, even more chromatically selective dots are now being constructed as a single layer, cut directly from recently available 3M tinted ScotchLite sheets. The color selectivity of these materials is good enough that we are also adding yellow and brown to our corral of recognized dot colors, extending *glimpser* accordingly.

5.2.3 voodoo

An application-independent geometric parsing toolkit called *voodoo* interprets the simple colored-dot-location output of the *glimpser* program. *voodoo* analyzes each unorganized per-frame collection of found color dots into a list of the unique patterns that have been registered with it by the application it serves. These patterns specify a sequence of colors; associated with each pair of adjacent color dots in a pattern is a distance, and with each contiguous triplet of dots an angle. These two parameters – the distance between each pair of dots and the angle through each triplet, along with the dots' color sequence – are enough to uniquely define any arbitrary pattern; as discussed earlier, we assign one such pattern to each of the client system's known objects, both physically (colored dots are affixed to the top of the object) and computationally (the pattern is registered with *voodoo*).



PATTERN 1:

color	dist	angle
red		
green	1.0	
green	1.0	0°

PATTERN 2:

color	dist	angle
red		
green	2.0	
green	0.5	0°

PATTERN 3:

color	dist	angle
red		
green	1.0	
green	1.0	120°

Fig 14: dot-pattern definitions in *voodoo*

An adjustable tolerance, definable for each distance and angle specification, permits *voodoo* to absorb the inevitable inaccuracies and occasional single-pixel indecisions of machine vision algorithms. This tolerance mechanism further makes possible the definition of 'parametric' patterns: patterns that are correctly recognized throughout a range of angles or spacings, but which may then use the particular value of angle or spacing to set some flexible parameter (the variable curvature of a lens, etc.) [13].

voodoo also provides an 'object persistence' scheme, so that when low-level vision fails for a frame or two – or when users' hands intermittently occlude dots – the affected objects exhibit a bit of 'temporal inertia'. Objects can thus continue to exist for a short while even in the absence of positive visual data.

5.2.4 dee-voodoo

voodoo in its 'solitary' guise is responsible for reporting to its master application the identities and locations of all objects seen by its I/O Bulb. Making the same report to every other I/O Bulb as well is one way of connecting together multiple such instances in the same room. Thus *dee-voodoo* ('distributed *voodoo*') is a set of extensions to the existing software – entirely transparent to the higher-level implementer – whose first task is to connect over a dedicated TCP/IP port to all other *dee-voodoo* processes that can be found in close network proximity. Each of the various resulting links is then used to effect a bidirectional transfer of geometry information: the initiating (newest) *dee-voodoo* describes its I/O Bulb's position, orientation, and associated surface dimensions, all with respect to some globally acknowledged reference, and receives in return the distant I/O Bulb's complementary particulars.

Following that preliminary exchange, every object recognized by the *dee-voodoo* process serving I/O Bulb **A** is not only reported to application **A**, but is also relayed to I/O Bulb **B**'s (and **C**'s and **D**'s and ...) *dee-voodoo* process, which in turn reports the object to application **B** *as if the object had been seen by I/O Bulb **B***. This entails a small preparatory step in which **A** first transforms the reported object's geometric description into **B**'s local coordinate system. The object-exchange is also of course reciprocated from **B** to **A**, and so on, in an ongoing relay mesh with $n^2 - n$ links.

Although we have at present only two I/O Bulbs, we have tested use of *dee-voodoo* to synchronize up to five independent application processes, using 'manual' mouse-and-keyboard manipulation for objects of the three bulb-less processes. Even with the consequent twenty point-to-point links in simultaneous operation, the participating systems evinced negligible lag (less than or equal to one frame-update time) between the movements of local objects and those of distant objects.

Naturally, an $O(n^2)$ communication scheme like this one is unlikely to scale very well. With this in mind we've recently begun to experiment with alternate connection topologies, including centralized (star-shaped) and annularly-distributed (shift-ring) approaches.

It is through the use of *dee-voodoo* that we finally begin to assemble individual I/O Bulbs – the atomic units of our graphical-display-and-capture world – into a true Luminous Room structure.

6 DESIGN FOR LUMINOUS ROOMS

If the visual aspect of computation to date is theatrical – the screen an empty proscenium whose contextless nothing can be filled with anything, and therefore filled with the luxury of very few inherent expectations – then visual design for the Luminous Room must be more like narrative filmmaking or cinema verite: graphical co-occupation of a world already filled with people, things, and assumptions. Thus: design that's about accommodation and cooperation.

The future of reactive, real-world graphics will surely have its own Rands and Tuftes, Leacocks and Gilliams. For now, however, we have identified a few general guidelines.

6.1 Physical Objects

A central characteristic of all the applications we've built for the Luminous Room thus far is their extensive use of physical objects. This is an element curiously absent from nearly all other graphics-in-the-real-world research systems, which tend rather to rely for input either on gestural or symbolic means (mediated by VR suits & gloves, 'Twiddlers', etc. – things that are part of the computer's world but fundamentally *not* part of the real world). A single notable exception is – again – Pierre Wellner's superb DigitalDesk project [14].

Objects, usefully, are an excellent way to represent complex state. The distribution of optics models in the *Illuminating Light* application, for example, itself contains a great deal of information (irrespective of the graphical/digital parts of the system). If we imagine implementing the same system, on the same physical tabletop, but without the optics models, and relying instead solely on 'hand-tracked' gestural input: it's clear that the optics must then be represented graphically – and that simply moving them around is suddenly much harder for the operator.

We find further that proper deployment of physical objects in Luminous Room interfaces manifests itself in the tight cognitive binding of these implements with the attending projective information. One facet of this physical/digital association is a strong sense of causality: *Illuminating Light*'s optics seem to act directly on the beams they modify. The resulting interactions thus extend the illusion of a tight causality (rotating a mirror causes the beam it's reflecting to sweep across the workspace) even though there is an *implementational* distinction between input and output.

6.2 Graphical Dynamism

Each of the five applications discussed at the paper's outset is marked by a constant graphical dynamism. Indeed, pains have been taken to incorporate subtle motion into every graphical construct that does not, by the nature of its content and meaning, demand stasis (shadows, for example, are obviously not free to dance around; but even laser beams, which clearly must not translate laterally, are represented by a dashed line that swims ever forward). We find that, as a general design principle for Luminous Room interactions, these small visual gestures are desirable for the following reasons:

- *Apparent life*. Slight ongoing motions reassure the Luminous Room occupant that the room is still respondingly alive. They also lend a modicum of personality to the application: not strictly necessary, but always welcome.
- *Disambiguation of the real and the virtual*. Early tests with largely static graphical systems showed that with fairly dense, interpenetrating collections of physical objects and digital projections, confusions could sometimes arise over the status of the projections. Slight motions of a sort unlikely to attend physical objects help to signal graphics' identity.
- *Increased resolution*. Because the resolutions at which our current Luminous Room applications operate (32 dpi down to 4 dpi) are significantly lower than those commonly provided by other displays, human parsing of text is often hampered. But – since these glyphs are anti-aliased – even sub-pixel motions can dramatically increase their comprehensibility. Text aside, the perceived resolution of all projected Luminous Room graphics is increased when these constructs are in motion.
- *Aesthetics*. If we understand the aesthetics of an interaction to be a function of clarity and detail, then the combination of the three effects just described certainly leads to a 'pleasanter experience'. More ineffably, applications that apply subtle motions to different parts of their graphical apparatus simply look better than those whose elements are static.

6.3 Boundarylessness

The history of computer-based display is a history of bounding rectangles: buttons, text-blocks, panels, frames, windows, desktops, and – ultimately – the CRT screen itself. Real-world graphics must be deployed with a significantly different philosophy. Painting a projector's-worth of rectangular frame onto the floor or wall makes the framed region a *window onto* something else; our goal, contrariwise, is to make projected graphical constructs *part of* the physical surround, to gracefully integrate these graphics with (possibly mobile) parts of the physical world.

Considered another way, the question is one of compositing. Screen-based pixels lie only over the blackness of the screen. But real-world pixels must be composited with reality; thus, as with any compositing task, the irrelevant but literal rectangular shape of the *ground* in which the intended *figure* pixels lie must be hidden by making all *ground* pixels transparent.

In the end, any boundaries must belong to the room itself: the edge of a table, the join of a wall with the floor, the moulding near the ceiling: for graphics in the Luminous Room, *these* are the borders that count and that must be respected.

6.4 Kinds of Applications

Four of the five applications described in this paper have subscribed to a remarkably similar characteristic: with the exception of the Chess-&-Bottle system, each has addressed a domain or phenomenon whose concerns are directly spatial in nature. Optical design is a question of the proper geometric arrangement of component elements; urban planning deals with the positioning and orienting of large architectural structures to solve both aesthetic and pragmatic problems; and so on. Generalizing from the current set of examples, and acknowledging that the Luminous Room's basic nature encourages interactions that involve physical objects arrayed and moved through space, we feel confident that a broad range of spatial applications will map naturally to similar workbench-style Luminous Room configurations.

While the Chess-&-Bottle application begins in a modest way to explore more abstract manipulations, understanding how to properly formulate non-spatial problems for treatment in a Luminous Room setting remains a longer-term challenge.

7 CONCLUSION

We have presented a broad overview of the Luminous Room, an infrastructure for distributing digitally-generated graphics and interaction throughout an architectural space. With one particular set of hardware and software techniques implementing the pervasive display and capture needs of the Luminous Room, and a handful of illustrative applications, we've described a small initial foray into an alternative graphical interaction field we believe to be extremely fertile.

ACKNOWLEDGEMENTS

Dean William Mitchell of MIT's School of Architecture and Planning, for unassailable enthusiasm for each part of the project and for the informed conviction that each'd be worth doing in the first place; Andy Dahley, Peter Underkoffler, and Paul Yarin, who collectively applied the steel-toed kick of design to occasional physical artifacts that really deserved it; Nicholas Negroponte; Gustavo Santos and Dan Chak, whose future we expect to transcend our mere present; and – most of all – Wendy Plesniak, who in 1996 catalyzed the entire project idea with a short sentence and whose influence has hovered over it ever since.

REFERENCES

[1] Agrawala, M., Beers, A., Frohlich, B., et. al. The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space, in *Proceedings of SIGGRAPH*, 1997, 327-332.

[2] Bolt, R. "Put-That-There": Voice and Gesture at the Graphics Interface. *Proceedings of SIGGRAPH*, 1980, 262-270.

[3] Cruz-Neira, C., Sandin, D., and DeFanti, T. Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE, *Proceedings of SIGGRAPH*, 1993, p. 193.

[4] Feiner S, MacIntyre B, Seligmann D. Knowledge-Based Augmented Reality. *Communications of the ACM*, Vol. 36, No. 7, July 1993

[5] Frisch, E., Hasslacher, B., and Pomeau, Y. Lattice-Gas Automata for the Navier-Stokes Equation. *Physical Review Letters*, 56, 1505-8

[6] Ishii, H. and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *Proceedings of CHI '97*, March 1997.

[7] Krueger, M. Artificial Reality II. Addison-Wesley, 1991.

[8] Jacobsen, J., Comiskey, B., et al. The Last Book. *IBM Systems Journal*, Vol. 36, No. 3, International Business Machines Corporation, 1997.

[9] Paradiso, J., and Hsiao, K.-Y. Swept-Frequency, Magnetically-Coupled Resonant Tags for Realtime, Continuous, Multiparameter Control. *Proceedings of CHI '99, Late Breaking Results*, May 1999.

[10] Pentland, A. Smart Rooms. *Scientific American*, v. 274 no. 4, pp. 68-76. April 1996.

[11] Raskar, R., Welch, G., Cutts, M., et al. The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays. *Proceedings of SIGGRAPH*, 1998, 179-188.

[12] Smith, J., White, T., Dodge, C., Allport, D., Paradiso, J., and Gershenfeld, N. Electric Field Sensing for Graphical Interfaces. To Appear in *Computer Graphics and Animation*, Special Issue on Novel Input Devices.

[13] Underkoffler, J. and Ishii, H. Illuminating Light: An Optical Design Tool with a Luminous-Tangible Interface. *Proceedings of CHI '98*, April 1998.

[14] Wellner, P. Interacting with Paper on the DigitalDesk. *Communications of the ACM*. Vol. 36, No. 7: 87-96, July 1993.