

Multi-Projectors and Implicit Interaction in Persuasive Public Displays

Dietz, P.; Raskar, R.; Booth, S.; van Baar, J.; Wittenburg, K.; Knep, B.

TR2004-021 March 2004

Abstract

Recent advances in computer video projection open up new possibilities for real-time interactive, persuasive displays. Now a display can continuously adapt to a viewer so as to maximize its effectiveness. However, by the very nature of persuasion, these displays must be both immersive and subtle. We have been working on technologies that support this application including multiprojector and implicit interaction techniques. These technologies have been used to create a series of interactive persuasive displays that are described.

To appear in AVI 2004

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Publication History:

1. First printing, TR-2004-021, March 2004



Multi-Projectors and Implicit Interaction in Persuasive Public Displays

Paul Dietz, Ramesh Raskar, Shane Booth, Jeroen van Baar, Kent Wittenburg, Brian Knep
Mitsubishi Electric Research Laboratories
201 Broadway
Cambridge, Massachusetts, USA
(+1)617-621-7500

{dietz,ramesh,booth,jeroen,wittenburg}@merl.com, bk@blep.com

ABSTRACT

Recent advances in computer video projection open up new possibilities for real-time interactive, persuasive displays. Now a display can continuously adapt to a viewer so as to maximize its effectiveness. However, by the very nature of persuasion, these displays must be both immersive and subtle. We have been working on technologies that support this application including multi-projector and implicit interaction techniques. These technologies have been used to create a series of interactive persuasive displays that are described.

Categories and Subject Descriptors

I.3.8 [Computer graphics]: Applications, J.4 [Social and Behavioral Sciences]: Psychology, Sociology, J.7 [Computers in Other Systems]: Consumer Products

General Terms

Algorithms, Design, Experimentation, Human Factors

Keywords

Implicit interaction, persuasive displays, ubiquitous computing, projective displays, multi-projectors, proximity detection, retail environments, proximity sensors

1. INTRODUCTION

Public displays can serve many purposes. They can inform, entertain, or attempt to persuade [6]. Although public displays that inform or entertain are common, persuasive displays represent the bulk of public displays. Commercial advertising is the predominant form. Billboards line our highways, and neon lights blaze bright in our cities.

Recent advances in computer video projection now allow inexpensive public displays to be created that can be used in areas with significant ambient lighting. Indeed, it is not unheard of to see a video projector being used in a retail display. However,

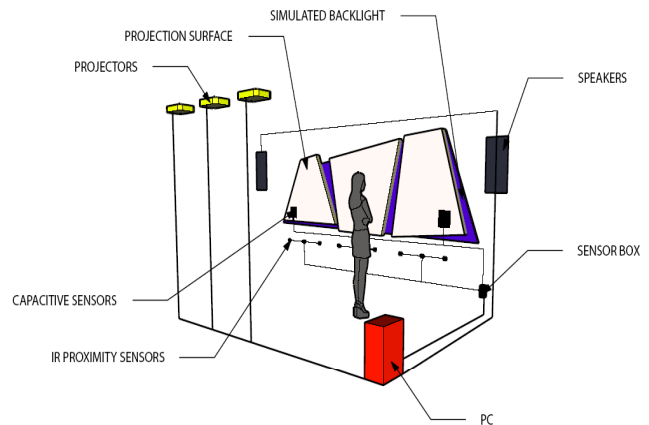


Figure 1 Overview of our advanced retail display. Three overlapping projectors are used to display the imagery. Sensors in the display can detect viewers at different distances, and implicitly change the displayed imagery.

most installations simply use the projector to create a large, flat display in a special display area.

While a video projector can be used to create “a big monitor,” such displays are not taking advantage of the unique properties of projection. It has been noted [17] that unlike any other display technology, projection decouples the size and shape of the device from the size and shape of the display. So a small projector can not only create a large display, it can project onto curved and other oddly shaped geometries. Because virtually any surface can be projected on, a space can be filled with imagery in a way that could never be achieved with CRT, LCD or plasma displays. This ability to immerse a viewer in media is extremely useful in persuasive applications—the viewer need not imagine the results of some desired behavior—they can actually see what it would be like.

While immersive media is a very powerful persuasion tool, it lacks universality. There is no one right presentation that appeals to everyone, all of the time. Thus there is a need for creating an interactive system that can adapt to the current viewers. Once again, projection is ideal because media can be changed at little cost.

In the field of human-computer interaction, we tend to think in terms of explicit actions taken by a user to consciously direct the computer to do something. However, in the case of persuasive displays, implicit interaction is the rule rather than the exception.

People do not generally wish to be persuaded to a viewpoint they do not currently hold. So the system must typically choose media without any explicit guidance from the viewer. Instead, the system attempts to do the right thing based upon *sensing* the actions of the viewer. Although the system is reacting to the viewer, the viewer is not in control. In fact, viewers should be unaware that it is their actions that are causing the display to change.

In this paper, we will describe our initial efforts to create implicit interactive persuasive displays and their supporting technologies. First, we shall begin with a brief discussion of related work and how our own thinking evolved on this issue. Then, we will examine projection techniques including nonplanar projection and multi-projector displays. Next, we describe their use with a series of retail displays and the role of implicit interaction. A diagram of our most advanced prototype is depicted in Figure 1.

2. BACKGROUND

Our interest in this area began with the observation that projectors might be useful tools in creating retail displays. After visiting a number of stores, it quickly became apparent that video systems are already extremely popular. We noted two frequent uses. First, TV/VCR or TV/DVD units were often used in small displays to act as tireless sales staff, constantly pitching a particular product, “infomercial” style. More rarely, a large display would be used for this purpose. The second, more subtle use, could best be described as “creating environment.” We found many examples where a series of monitors were used to show music videos without any apparent connection to the products for sale. Certainly, projectors could also be used for either of these purposes.

Another trend that we noted was the use of theatrical lighting techniques to highlight products. In most higher end stores we visited, the lighting systems were quite complex, and clearly a great deal of thought had gone into their design. But beyond static systems, we found a number of retailers were using computer-controlled lights to create various effects ranging from simple color changes to fairly spectacular light shows. Once again, it occurred to us that much of the functionality of the expensive theatrical lighting equipment that we saw could be inexpensively emulated with video projectors.

The uses of theatrical lighting and standard video displays are fundamentally different. As a rule, the video displays showed their images in well-defined, flat, separate areas, distinct from the merchandise. The lighting systems did just the opposite. They generally focused directly on the merchandise itself. Projectors have the advantage that besides just being a display, the projective nature of displaying its image enables us to augment our surroundings, e.g. objects, surfaces, etc., with imagery. Lighting and video display could thus be done simultaneously by the projector.

Our first projector-based retail display began as an effort to emulate two functions: “pitch” a product, and create an environment. However, before describing our display, we first discuss the underlying projection technology.

2.1 Previous Work on Projected Displays

At MERL, we have a diverse research program examining projected displays. Two of the major areas under investigation have been the use of projectors in applying color and texture to

general, i.e. *nonplanar*, objects (shader lamps), and the automatic alignment of multi-projector displays on both planar and non-planar surfaces and objects.

Shader Lamps

Projectors can be used to change the surface appearance of neutral colored physical models by illuminating them with rendered images [20]. We use digital light projectors and computer graphics to form *Shader Lamps* that effectively reproduce or synthesize various surface attributes statically, dynamically, or interactively. While the results are theoretically equivalent for only a limited class of surfaces and attributes, our experience is that they are quite realistic and compelling for a broad range of applications. In this type of projector-based augmentation, complete illumination of complex 3D shapes is made possible by rendering geometric models of those shapes from multiple overlapping projectors. The geometric models are pre-authored with detailed color, texture, and material properties to reproduce desired surface appearance.

The idea of using projectors to graphically animate physical objects in the real world has been used by some to compelling effects. Based on the underlying complexity, the previous work in using projectors for augmented reality can be divided into two main groups: (1) projection of useful information on a planar surface, and (2) more complex insertion of 3D shapes and attributes into the real world. 2D interactive images can be projected onto flat surfaces in a room to enhance the user's environment. Examples include Luminous room [29], digital desk [31], and smart whiteboards. 3D augmentation examples include ShaderLamps [20] and the related project Being There [23].

Naimark [14] used a rotating movie camera to film a living room, replete with furniture and people. The room and furniture were then painted white (neutral), and the captured imagery was projected back onto the walls using a rotating projector that was precisely registered with the original camera. This crucial co-location of the capturing and displaying devices is common to most of the current demonstrations that use pre-recorded images or image-sequences. A limited but compelling example of this idea is the projection of pre-recorded video to animate four neutral busts of singing men in Walt Disney World's "Haunted Mansion." In addition, a patented projector and fiber optic setup animates the head of the fictional fortuneteller "Madame Leota" inside a real crystal ball [11].

Slides of modified photographs augmented with fine details are also used with very bright projectors to render imagery on a very large architectural scale. A well-known modern realization of this idea is the Son et Lumiere light show on the Blois castle in the Loire Valley, France

All these systems create compelling visualizations. However, the cumbersome alignment process can take several hours even for a single projector. Our technique avoids this problem by forming a 3D geometric understanding using well-known computer vision techniques and then moves beyond simple image projection to reproduce novel appearance.

Multi-Projector Displays

Large projector arrays are popular because they allow a practical solution to generate high-resolution and bright images by tiling together a set of projectors. Older systems, such as video walls,

use a two dimensional matrix of rear projection with abutting image edges separated by a small but visible gap. As shown in Figure 2, newer systems [4][27] use overlapping projectors with facility for precise manual electro-mechanical adjustment for image registration and cross-fading. Due to such design constraints, the installation and operation of such systems is extremely expensive. The installation of these displays is quite tedious, requiring precise projector overlap, often needing orthogonal projection to the screen. This arguably is the most prominent drawback of large format display design. Research into techniques for automating this registration process is helping alleviate this time consuming setup [21][22][17][3].

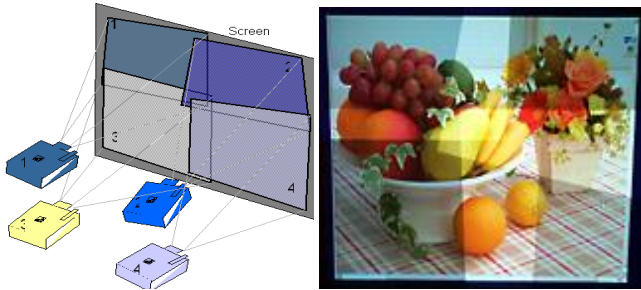


Figure 2 Projector mosaic creates a seamless display

Our goal is to significantly reduce the support and infrastructure cost by providing a flexible calibration and rendering technique that can adapt to a given projector array configuration. Further, our goal is to reduce the time required for setup and registration to a few seconds. The calibration technique is designed to be very simple to use and free of any human interactions.

This approach to displays allows very wide aspect ratios, short throw distance between projectors and the display surfaces, hence higher pixel resolution and brightness, and the ability to use heterogeneous units. We will now describe the series of retail displays we built using the concepts discussed in this section.

3. A FIRST RETAIL DISPLAY

Our first prototype persuasive display simulated a retail environment [32]. It is shown schematically in Figure 3. There were three distinct projection areas that were designed to accommodate the geometry of the room. These included a car display, a shoe (interactive shelves) display and an environmental display. Obviously, we made no attempt to keep the themes consistent from display to display. We wished to show the broad applicability of our techniques instead.

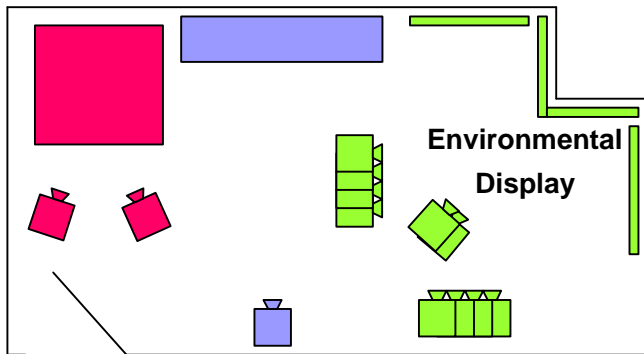


Figure 3 Diagram of the original retail display

3.1 The Car Display

The car display created the illusion of a small car driving down a road (Figure 4) and has been well described in [20]. We imagined that this piece of our display represented part of an automobile showroom that allowed customers to see a car and its various color options. Unlike a typical panel of color swatches, Shader Lamp technology allows us to create a far more engaging display. The emotional appeal of hearing and seeing a car in 3-D, in your preferred color, driving around, is very powerful. Unlike most other 3-D technologies, Shader Lamps provide a real 3-D image that many people can enjoy simultaneously without the use of special glasses. They can even touch it if they want. From the customer viewpoint, it makes sense to look at this display to see the color options. But from the car dealer's perspective, that is just an excuse to draw the customer into a much deeper emotional experience. Many of the most effective persuasive displays use this technique—the desired information pulls the viewer into an emotional appeal.

In order to create this display, we began with a detailed 3-D model of the toy car that was painted diffuse white. The car was placed on a computer-controlled turntable and illuminated by two video projectors. After an alignment procedure, the content was warped in software to display correctly when projected onto the three dimensional surface. Speakers were also hidden in the display to provide realistic driving noises.

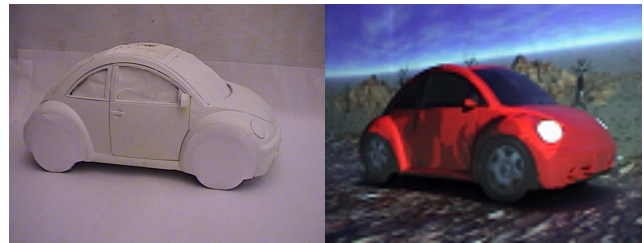


Figure 4 Model car without (left) and with (right) projection

The visual and audio content for this display included a number of different modes. Because the warping is calculated in real-time, all attributes such as color, lighting, background, animation, etc. can all be changed in real-time. Visually rotating the wheels and moving the background created the appearance of motion. A cartoonish appearance was achieved by rendering in cartoon style and projecting oval patterns on the wheels during moments of high acceleration or deceleration. The addition of professionally composed background music also added to the appeal and the effectiveness of the display.

3.2 The Environmental Display

Immersive displays are an extremely effective tool for persuasion. Many retailers take advantage of this fact and fill their stores with props and imagery to create a sales conducive environment. For example, winter scenes evoke memories of the cold and help sell coats even on warm fall days. Retailers go to tremendous lengths to create environmental displays. Unfortunately, like any constant stimulus, if the display remains unchanged, viewers become blind to it. Thus, it is critical for stores to constantly change displays before they lose their effectiveness. New graphics need to be manufactured, distributed and installed on a regular basis. This is tedious and expensive.

By contrast, video content is easily created, distributed and “installed.” Many retailers have already realized this, and are using large numbers of video displays in their shops to create atmosphere. Unfortunately, traditional video display devices are physically small compared to most stores, so the ability to create an immersive display is limited.

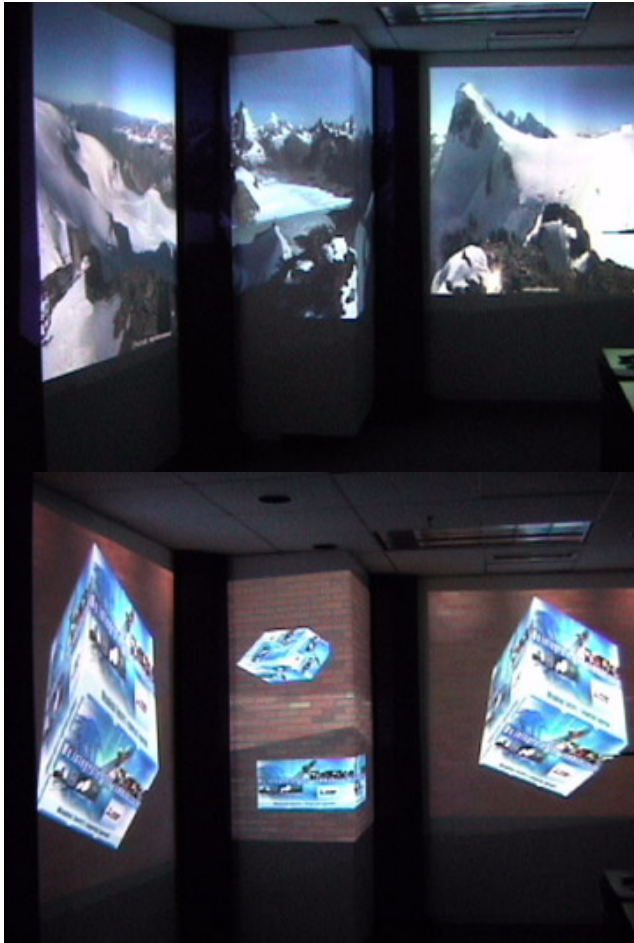


Figure 5 (a) An evocative scene being displayed with multiple projectors across a nonplanar surface (b) A change of that scene showing brick wall texture and a synchronized 3D display of a rotating cube

Although the display from a video projector easily dwarfs the largest conventional displays, it is still small compared to the scale of a large retail store. In order to fill such large spaces, multiple projectors will be required. We have developed techniques that allow us to automatically stitch projected displays together to create giant, seamless displays. Environmental displays such as those in Figure 5 demonstrate how a multi-projector system can create a dramatically more immersive display. The figure shows two examples of content displayed on our large area seamless display. This system uses 10 projectors that warp the content to the geometry of the room.

To create large area displays using projectors in retail environments poses several challenges. We have chosen to use front projection for two important reasons. First, it is much simpler to install than rear-projection systems because it does not

require additional space for the installation of projectors behind the walls nor does it require installation of large transparent screens. Second, front projection allows us to project directly on existing architectural features and products to enhance their appearance.

One challenge with front-projection is that customers could interfere with the projectors’ light path and cause shadows on the display that would greatly reduce the visual impact. Rehg et al. [24] and Jaynes et al. [7] have introduced techniques for dynamic shadow removal for front projection displays. These techniques require a high computational load, especially if many customers at once could be interfering with the projectors’ light path. Together with the computational load of displaying multi-media content, dynamic shadow removal techniques are not yet practical for retail environments. To overcome the challenge of customers interfering with the display we have mounted multiple projectors on the ceiling and project in overlapping fashion onto the walls under acute angles with short throw distances. Such non-orthogonal projections will result in keystone distortion, which needs to be addressed in order to project undistorted contents. For planar surfaces we have developed an automatic keystone distortion technique using a camera [18][19].

For typical retail environments the walls are not necessarily large planar surfaces but could be interrupted by facets with different orientations, e.g., partially embedded support columns, extrusions, and corners between walls. To register multiple projectors on non-planar surfaces a geometric model of the underlying surface is required. Registration between the physical surface and the projections can then be achieved by matching pre-determined points on the surface with points in the projector image [20]. A requirement to facilitate easy deployment of multi-projector systems like these is to make projector-surface registration as simple as possible. We have pursued a number of techniques, including manual registration by clicking with the mouse pointer on key points as they are projected onto the environment, camera based automatic calibration, and newer techniques involving fiber optic sensors that interpret structured light projected with the projectors. Although the last method would require a one-time installation of embedded optic sensors in the wall, the ease-of-use with respect to registration compensates for the installation requirement.

Finally, the lighting in retail environments could prove a challenge for projection displays. In many retail environments the lighting is a carefully controlled aspect of the atmosphere of the environment. Instead of compensating for the lighting environment in the projection, the large area displays could complement the lighting environment in retail spaces. We can even imagine a future where the majority of a store’s lighting comes from video projectors. For the lighting designer, it is the ultimate programmable light.

The prototype shown in Figure 5 yielded a vision of how large multi-projector displays could be embedded in public spaces. However, the prototype suffered from the fact that the hardware and software requirements for the multi-projector displays were daunting. In this version, each projector required an accompanying CPU, as well as an additional CPU to coordinate the display of content across the CPU/projector pairs.

3.3 The Interactive Shelves Display

The shoe display began as our attempt to demonstrate how a projector could be used to replace various monitors and theatrical lights currently used in displays. We soon realized that this was also an ideal platform to test some of our ideas about interactive displays, and we rapidly went through a number of iterations of the system.

The original shoe display consisted of a single projector displaying on four shelves. Although we made no attempt to model the 3-D geometry for this display, we did design the 2-D content specifically for this arrangement. The content was created in Macromedia Flash and included a specially composed audio track. Figure 6 shows the setup, including the sensors that were added later.

Unlike our other displays, all of the content for this display was created by professional designers. This was a rather humbling experience for the engineers involved with the project because it dramatically made the point that, ultimately, the content drives everything. A beautiful sign will be far more effective than the most sophisticated interactive 3-D display if the content is poor [8]. We resolved that design would be the focus of further efforts, and that technology would be created to meet specific design needs rather than the other way around.

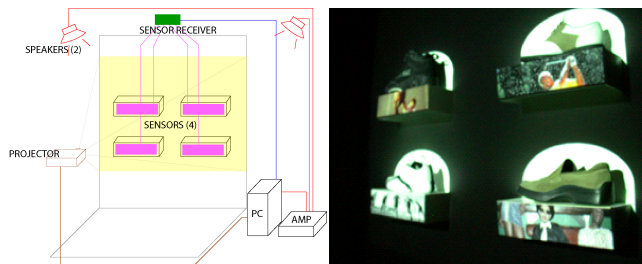


Figure 6 The shoe display setup (left), and the shoe display in action (right)

A projector is essentially a blank canvas for designers. In the shoe display, we simulated many different kinds of theatrical lights and video screens (many of different shapes and sizes) all coming from the same projector. Four full motion video screens appear on the protruding shelves, each with a resolution that compares well with a traditional monitor. At times these would be replaced with spotlights that highlighted the shoes mimicking theatrical effects. Large graphics were located above the shoes. Some of this is visible in Figure 6. Everything was timed to a musical score that greatly enhanced the emotional connection made with the viewers.

Although this display was audio-visually stunning, it suffered from scale issues. From a distance, it was very attractive. But, when a viewer came closer to inspect a shoe, the display bordered on distracting. It became clear that the display needed to respond differently depending on how close the viewers were. Thus began our investigation into implicit interfaces for persuasive displays.

The shoe display was modified to include proximity sensors in the shelves. This involved lining the shelves with aluminum foil and connecting them to a sensitive capacitive measurement circuit based on the Maxwell method of intermittent currents [13]. A four channel system was constructed to detect human presence up to ~10cm away from the shelf surfaces. This is far cheaper and more robust than vision-based techniques.

The software system also needed modification to support the new sensors. We decided to use Visual Basic to read the sensors and control an ActiveX Flash player. This turned out to be a simple and effective combination.

New interactive content was created to take advantage of the sensors. After the initial multimedia movie played (30 seconds long) a projected image appeared that said, "Please choose the shoe of your choice." As a customer's hand approached a shelf, content would be triggered that would talk about the particular shoe that they were interacting with. We created content for each individual shoe that we had on the shelves. For example, when a viewer touched the fashion shoe, they were immediately surrounded with imagery depicting hipsters wearing that particular shoe, accompanied by loud drum and base music.

This pattern was repeated for each of the shoe types—running, basketball, fashion, and golf. Each had its own content and musical score. Not only had we solved our scale problem, but also by detecting the shoe of interest, we were providing targeted content at just the right moment. We made grandiose comparisons to Internet shopping and talked about bringing point and click functionality back to the physical world.

While most people seemed to enjoy playing with the display, we soon started to see the inherent flaws with this approach. Viewers were not as interested in the informational content as we had expected. Few people actually read the text information. Worse yet, some viewers did not like the loud music that was associated with each shoe. They said that it made them feel singled out among the test subjects and they didn't want to be the center of attention while browsing.

In analyzing these reactions, we slowly began to realize what had gone wrong. We had presumed that the interactions should be explicit. Users should consciously take action to causally affect the display. But this really was a gross departure from the philosophy behind the other displays in the room. Salesmanship is a subtle art. Another problem with this display was that it was inherently single user—only one informational display could be displayed at a time because each took up all the space above the shoes.

For our next generation shoe display, we addressed these issues. First, the content was modified to subtly draw in a customer. The loud blaring music and rapid change graphics were replaced by content that subtly shifted and changed when a customer interacted with the shoe on the display. Another major improvement was the ability to accommodate multiple users interacting with the display. The more subtle content shifts still enhanced the experience for each user without distracting other viewers.

Our new model of how this interface should work became that of a "magical" experience. Customers do not like having people try to sell them things. This may be a result of over-saturation of advertising on all kinds of media. The key to content is to subtly pitch the product to the consumer, analyze the behavior of the consumer, and then proceed accordingly. The model that we are trying to achieve is that of a good salesman. As a customer, you really should not even know that you are having a product pitched to you. The content of the display will be "coincidentally" appealing to you and appropriate to your physical position with

respect to the merchandise Customers are not normally conscious of the causal relationship between their actions and changes in the display. We refer to this as implicit interaction.

4. A SECOND RETAIL DISPLAY

We next decided to create a more ambitious display that mixed implicit interaction techniques with multi-projector systems. In addition, we wanted to address two other barriers to use: cost and content creation tools.

The theme for this new display is audio products for personal use. Figure 7 shows the setup for the new retail display. (See also Figure 1.) There are three focus areas in the display, each with its own interactions. The entire display is illuminated by three video projectors with the imagery suitably blended and warped to create a single, long, high-resolution display. All three projectors are connected to a single PC with a three-headed graphics board. In contrast to our earlier multi-projector displays, we now require only one PC for the entire setup.

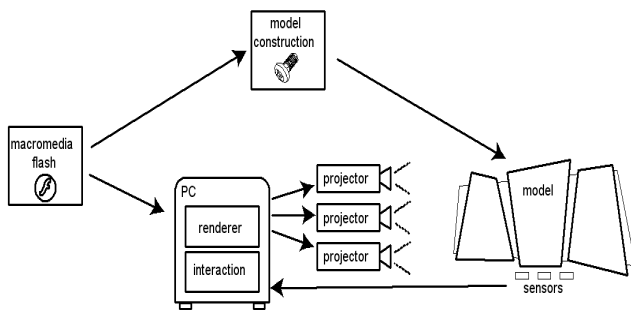


Figure 7 A schema of the system for our second-generation prototype

In addition to capacitive proximity sensors, this display also features infrared distance sensors that have a range of ~1.5 meters, allowing it to detect people approaching the display from a substantial distance. The capacitive and infra-red sensors are connected to the PC via a serial port.

4.1 Rendering

Currently, video walls in retail applications typically scale NTSC content up through various signal-processing methods. But for our application, people get very close to the display. Therefore, it is critical to be able to display full resolution content on our multi-projector display. Here we consider issues in approaches to rendering and associated authoring tools necessary for our goal.

A player that plays pre-rendered movies is a general solution. Authoring it boils down to specifying which movie to play and when to play it based on sensor data. The advantage of this method is its independence from any one graphical authoring tool. There are, unfortunately, several drawbacks.

First, pre-rendered movies must be warped to align them to the projectors. Every time we recalibrate the projectors, we have to re-warp the movies. This slows down the calibration process and introduces filtering artifacts.

Second, since the movies are pre-rendered, we have lost their internal structure—all we have are pixels. Therefore, we cannot

change the movies on the fly to adjust for customer or environmental changes. In other words, we cannot make a very dynamic system. We can write a system that transforms movies, plays one on top of another, fades between them, and renders other effects, but then we would have to write our own in-house authoring tool.

Third, we run into hardware limitations in that we cannot push pixels onto the screen fast enough. For a horizontal, three-projector system with 1024-by-768 projectors, 5% overlap, and a 5% border, we would optimally have 2867-by-691, or 1981097, pixels. At four bytes per pixel, this gives us about 7.5 megabytes of data per frame, or, at 30 frames per second, 225 megabytes per second—a prohibitive data rate for one PC to handle.

Due to these drawbacks we decided to implement a custom player based on Macromedia's Flash and rendering the contents using OpenGL. Macromedia Flash has an open file format, called SWF, and several open-source projects are available that read, write, and even partially render this format. This gave us the confidence that we could write our own player, leveraging off of open-source code when possible

More importantly for us, Flash is a 2D vector-based tool. It can display images and movies, but its focus is objects—including strokes, fills, and splines—rather than pixels. These object descriptions are more compact representations than pixels, and can be scaled arbitrarily without a loss of quality. Arbitrary scaling is important for the high-resolution displays we are aiming to create. Compact representations are important for reducing bandwidth—sending a polygon's outline to a graphics board takes far fewer bits than sending the pixels from a rasterized polygon.

Additional advantages of writing a custom player that renders with OpenGL are (i) we can offload the rasterization to the graphics card, allowing us to utilize the powerful performance of graphics cards nowadays and free the CPU for other tasks, and (ii) incorporate a single pass warping in the rendering.

4.2 Multi-Projector Display

The shoe display introduced the idea of multiple screens. In that display, the background screen is a large rectangular plane, and each of the shoes sits on a platform that doubles as a miniature screen. The audio products display extends this idea, using three large foreground screens and one large background screen. Although we have created technologies for arbitrary projection surface geometries, creating content for these systems is inherently complex. In an effort to avoid 3-D modeling, we decided to stick with planar screens, but with funky quadrilateral shapes.

Like the shoe display, the foreground and background screens are at different depths. However, with a multi-projector display, this raises the difficult issue of blending and warping at different depths.

The foreground screens are prominent, and they show the most important content. Therefore, we decided to focus our alignment efforts on these screens. We have three projectors, each shining upon a single screen. The projectors do not overlap on the screens, but do overlap on the background area between the screens. Since the background is where we placed our

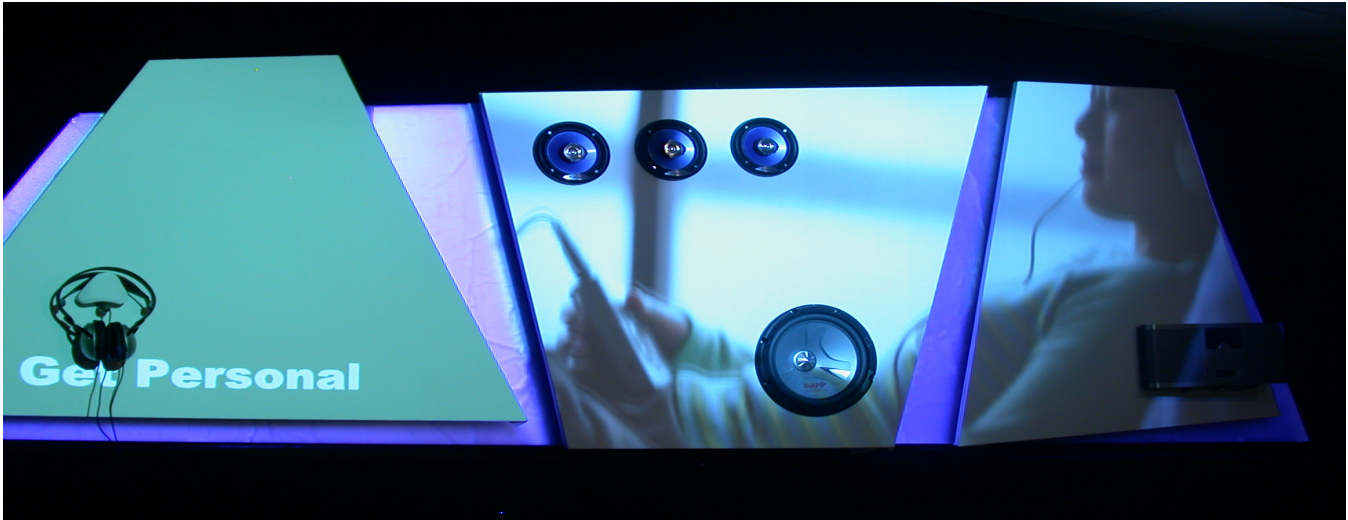


Figure 8 The second generation retail display prototype. The display consists of three non-rectangular focus areas with a translucent background in between. The projection of multiple overlapping projectors is blended to create a large display.

translucent/reflective surface, focus and alignment are not as important in these areas. This approach greatly simplified our inter-projector alignment problems.

To keep the images aligned with the physical model, each projection can be warped by a homography. The homography is specified by four input points and four output points—it warps each input point to the corresponding output point. To make the input points meaningful, we use the corners of the foreground screens, and we get these corners from geometry encoded in the Flash file.

To ensure that each projector projects onto only one foreground screen, we mask the projected image with a grayscale alpha-mask. In the overlapping regions, we blur the alpha mask so that the overlapped images blend. Currently these alpha masks are created by hand from Flash geometry, but the process could be automated.

Although the foreground screens are diffuse projection surfaces, we enhanced the visual appearance of the entire display with a back-lighting effect by layering a translucent plastic sheet over a metallic, reflective surface, with about an inch of space between. The projected light diffuses through the plastic, reflects off of the backing, and diffuses again through the plastic back toward the viewer. This shows that even though we use front-projection, we are able to create the appearance of giving a backlit, glowing effect.

4.3 Implicit Interaction

To detect customers, each screen has an array of infrared and capacitive sensors. The infrared sensors measure the distance to the closest customer, and the capacitive sensors measure whether someone is touching or very close to a specific part of the screen—in our case, a product attached to the screen.

The new display has several interactive modes per screen, and allows multiple simultaneous interactions. The display switches modes according to what we believe could be a typical scenario of a customer’s behavior in a store. In such a scenario we identified four modes: idle mode, attract mode, approach mode, and pitch

mode. All modes, except the first, can be triggered independently on each screen.

Idle mode plays when no viewers are nearby, which would be the case when the customer first enters the store, or the customer looks at the merchandise from a relatively large distance. In the case of our prototype display the content is an animation loop showcasing the products. The display will switch to attract mode once a viewer comes within the infrared sensor’s range. This mode is designed to implicitly draw viewers into the display. Contents displayed in this mode could for example, depict typical usages of the product (Figure 8), or create an environment, as discussed earlier in section 3.2.

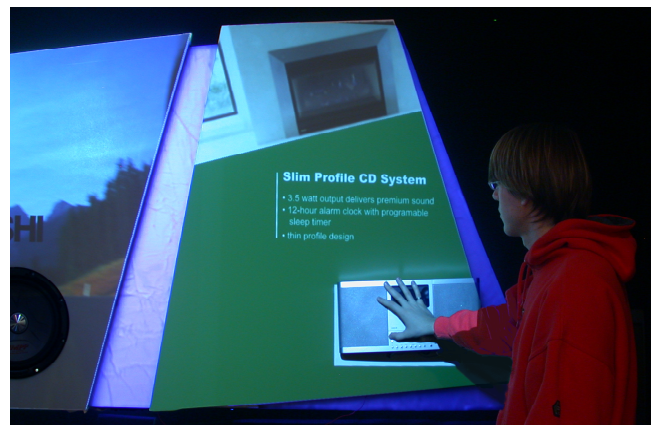


Figure 9 The display implicitly provides the viewer with additional information about the product in close proximity

If a customer is interested in the merchandise and approaches the display, the display switches to approach mode. In this mode the display could present the user with advantages of using this particular product, or some general facts. In our prototype we chose to display single-word statements about the product. When a capacitive sensor indicates that the customer is very close to or touching the product, the display switches to pitch mode. Figure 9 shows our prototype display in this mode. The contents displayed in pitch mode are typically text and possibly additional images or

animations. In our prototype the display shows a few sentences reiterating key features, advantages, and usages.

Switching modes without the explicit triggering by the viewer allows the display to become a natural part of the infrastructure, “disappearing” into the retail environment.

5. RELATED WORK

Passive electronic signage has appeared in real-world environments from sports stadiums to airports to flagship retail stores and outdoor areas. Dynamic advertising for plasma-based displays can now be triggered by external events customizable by a display controller. It is available on a commercial basis [5]. However, the use of embedded front-projection, particularly with multiple projectors and geometric modeling, has yet to be adopted commercially.

A number of researchers have been investigating sensor-based interaction with large displays. For instance, Paradiso et al., have built a number of interactive installations that make use of sensors and trackers to enable interaction [15][16]. In one case, a commercial installation of TapTracker, a technology for sensing tapping on glass, led to measurable increases in retail store traffic after being installed in a shop window [16]. A sample of other work on large interactive wall and table displays is available from a 2002 UbiComp workshop [28].

Probably the work closest to that described here has been done at IBM Research in a multi-year project based on the Steerable Display [10]. The Steerable Display uses a single front-projector with a controllable mirror to enable displays to be dynamically steered to different locations within a single room environment. Vision-based interaction is enabled with a camera and software that interprets various types of hand gestures as well as tracks an individual’s location in the space. Recently the group has also been investigating the application of such technology to retail environments [25][26]. Some of group’s philosophy is similar to our own. For instance, a goal is to make information “appear as naturally... as possible.” It should not interfere with the normal shopping experience, but rather enhance it. The group has come up with a prototype for interactive shelves as we have. However, their approach tends to be more informational than the one we advocate. For instance, they created informational displays for clothing shelves that are triggered by a shopper putting a hand in a particular bin [25]. They included touch icons that controlled information about sizes and available stock. As the authors candidly admit, an evaluation walkthrough brought out significant design problems. For instance, their assumption that users understood the relationship between their actions and the information display turned out not to be valid much of the time. We hope to have circumvented these issues with our model for implicit interactions. Also, our sensing technology is less expensive and more robust than that used by the IBM group and is less prone to delay and recognition failure.

It is worth noting the relationship of our work to Weiser’s well-known treatise on ubiquitous computing [30] that has had such an influence on the research community. In that treatise he argued that computers should “vanish into the background,” that the human-computer interface should “disappear,” i.e., require no conscious effort on the part of users. However, as Bellotti et al. point out, novel sensing interfaces face significant design

challenges if the goal is to have users intentionally and explicitly control the interaction [2]. For instance, how does a user know that a system understands it is being addressed? How does a user know that a system understands a command and is correctly executing the intended action? Widely known GUI conventions have provided solutions to such problems. Environmentally-embedded displays with sensors have not. We suggest here that persuasive displays need not, and perhaps should not, support users’ conscious intentions and explicit directives. Our model for implicit interaction does quite literally entail that the interface disappears from the user’s consciousness. The ethics of utilizing such systems are of course a real issue, as is the case with any advertising or other persuasive system, and will need to be addressed by practitioners [6].

6. RESULTS AND FUTURE WORK

Video projection is an ideal medium for creating very large scale, immersive imagery. These systems can create emotionally engaging experiences, making them ideal for persuasive applications. By their nature, such displays must be subtle, leading us to explore implicit interaction techniques.

Our assessments so far have been based on informal walkthroughs as well as on reactions to our demos at our company and also at retail trade shows. The problems of evaluation for ubiquitous computing systems such as ours have been articulated in [1]. We agree with the view that substantive evaluation for a system such as ours requires real use of a system deployed in an authentic setting. We are actively seeking a retail partner to host a test display so that we may evaluate the effectiveness of our methods under real-world conditions.

Although we have specifically focused on implicit persuasive displays in retail environments, one could envision other application areas as well. For example, implicit displays could enhance the user experience in museums, providing additional information about the particular object, or even highlight areas of significance. We note that in an application area like this, the display need not necessarily be persuasive.

This paper has focused on projectors as real-time output devices, and we have introduced our vision of implicit interaction. Our next step is to integrate far more capable sensors along with real-time modeling of user tastes. This will allow our displays to automatically discover useful correlations, and drive the displays in a more optimal fashion. Although use of vision and other advanced sensing technologies can cause privacy concerns [9], we hope to avoid these by performing all preference predictions without explicit viewer identification.

7. ACKNOWLEDGMENTS

This paper draws on the work of many people at MERL over the course on several years. While space prevents us from listing them all, special thanks must go to Remo Ziegler, Paul Beardsley, Johnny Lee, Dan Maynes-Aminzade, and Thomas Willwacher for their fundamental efforts in this area. We also thank Henry Kaufman for his contributions to the audio products display. Peter Wilder, who composed and performed most of the music for our displays, is to be credited for much of their emotional impact.

8. REFERENCES

- [1] Abowd, G., Mynatt, E. "Charting Past, Present, and Future Research in Ubiquitous Computing," *ACM Transactions on Computer Human Interaction* 7(1), 29-58, 2000.
- [2] Bellotti, V., Back, M., Edwards, W.K., Grinter, R. I., Henderson, A., and Lopes, C., "Making Sense of Sensing Systems: Five Questions for Designers and Researchers," *Proceedings of CHI '02*, pp. 415-422.
- [3] Y. Chen, D. Clark, A. Finkelstein, T. Housel, and K. Li, "Automatic alignment of high-resolution multi-projector display using an uncalibrated camera," In *Proceedings IEEE Visualization 2000*.
- [4] ComView Visual Systems. Comview viewscreen and viewmaestro product sheets. <http://www.comviewvs.com>.
- [5] CoolSign, <http://www.coolsign.com>.
- [6] Fogg, B.J., *Persuasive Technology: Using Computers to Change What we Think and Do*, Morgan Kaufmann, 2002.
- [7] Jaynes, C., Webb, S., Steele, R. M., Brown, M., Seales, W. B., "Dynamic shadow removal from front projection displays," in *Proceedings of IEEE Visualization, 2001*.
- [8] Khaslavsky, J., Shedroff, N., "Understanding the Seductive Experience," *Communications of the ACM* 42(5), May, 1999, pp. 45-49.
- [9] Krill, P. "Attention Shoppers: You May Be Studied," *CNN.COM/SCI-TECH*, January 15, 2002, <http://www.cnn.com/2002/TECH/internet/01/15/shopping.studies.id/>
- [10] Levas, A., Pinhanez, C., Pingali, G., Kjeldsen, R., Podlaseck, M., Sukaviriya, N., "An Architecture and Framework for Steerable Interface Systems," in *Proceedings of UbiComp 2003*, LCNS 2864, pp. 333-348, Springer-Verlag, 2003.
- [11] G. E. Liljegren and E. L. Foster. Figure with Back Projected Image Using Fiber Optics. US Patent # 4,978,216, Walt Disney Company, USA, December 1990.
- [12] K. Low, G. Welch, A. Lastra, H. Fuchs. Life-Sized Projector-Based Dioramas: Spatially Real and Visually Virtual. Sketches & Applications, SIGGRAPH 2001.
- [13] James Clerk Maxwell, "775. Intermittent Current" in *A Treatise on Electricity and Magnetism, Vol. II, 3rd Edition*, Oxford, 1892.
- [14] M. Naimark. Displacements. An exhibit at the San Francisco Museum of Modern Art, San Francisco, CA (USA), 1984.
- [15] Paradiso, J., "The Brain Opera Technology: New Instruments and Gestural Sensors for Musical Interaction and Performance," *Journal of New Music Research* 28(2), 1999, pp. 130-149.
- [16] Paradiso, J. "Tracking Content and Free Gesture Across Large Interactive Surfaces," *Communications of the ACM* 46(7), July, 2003, pp. 63-68.
- [17] Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S. and Forlines, C., "iLamps: Geometrically Aware and Self-Configuring Projectors", *Proceedings of SIGGRAPH 2003*, *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 22, Issue 3, pp. 809-818, July 2003.
- [18] Raskar, R., Beardsley, P., "A Self-Correcting Projector", in *Proceedings of CVPR: IEEE Conference on Computer Vision and Pattern Recognition (Kauai HI, December 2001)*, IEEE.
- [19] Raskar, R., van Baar, J., Chai, J.X., "A Low-Cost Projector Mosaic with Fast Registration", in *Proceedings of Asian Conference on Computer Vision (ACCV)*, January 2002.
- [20] Raskar, R., Ziegler, R., Willwacher, T., "Cartoon Dioramas in Motion", in *Proceedings of International Symposium on Non-Photorealistic Animation and Rendering (Annecy France, June 2000)*.
- [21] R. Raskar, M. Brown, R. Yang, W. Chen, G. Welch, H. Towles, B. Seales, and H. Fuchs, "Multi-projector displays using camera-based registration," In *Proceedings IEEE Visualization, 1999*.
- [22] Raskar, R., Immersive planar display using roughly aligned projector. In *IEEE Virtual Reality*, March 2000
- [23] R. Raskar, G. Welch, K. Low, D. Bandyopadhyay. Shader Lamps, Animating Real Objects with Image Based Illumination. Proceedings of the 12th Eurographics Workshop on Rendering, June 2001.
- [24] Rehg, J., Flagg, M., Cham, T., Sukthankar, R., and Sukthankar, G., "Projected light displays using visual feedback," in *Proceedings of the International Conference on Control, Automation, Robotics, and Vision (ICARCV)*, 2002.
- [25] Sukaviriya, N., Podlaseck, M., Kjeldsen, R., Levas, A., Pingali, G., Pinhanez, C., "Embedding Interactions in a Retail Store Environment: The Design and Lessons Learned," *Proceedings of Interact '03*, Zurich, Switzerland, September 2003.
- [26] Sukaviriya, N., Podlaseck, M., Kjeldsen, R., Levas, A., Pingali, G., Pinhanez, C., "Augmenting a Retail Environment Using Steerable Interactive Displays," in *Proceedings of CHI 2003 Extended Abstracts*, pp. 978-979.
- [27] Trimension Systems. <http://www.trimension-inc.com>.
- [28] UbiComp Workshop on Collaboration with Interactive Walls and Tables, 2002, <http://ipsi.fraunhofer.de/ambiente/collabtablewalls/papers/>.
- [29] J. Underkoffler. A View From the Luminous Room. Springer-Verlag London Ltd., Personal Technologies (1997) 1:49-59.
- [30] Weiser, M. "The Computer for the 21st Century," *Scientific American* 265(3), pp. 94-100, 1999.
- [31] P. Wellner, B. Interacting with paper on the DigitalDesk. *Communications of the ACM*, 36(7):87-96, July 1993.
- [32] Wittenburg, K., "Research on Public, Community, and Situated Displays at MERL Cambridge," Position paper presented at the Workshop on Public, Community and Situated Displays at ACM 2002 Conference on Computer Supported Cooperative Work, November 2002, Tech Report TR2002-45, <http://www.merl.com/papers/TR2002-45/>
- [33] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. Brown, "PixelFlex: A reconfigurable multi-projector display system," In *Proceedings IEEE Visualization, 2001*.