PROJECTION – VISION SYSTEMS: Towards a Human-Centric Taxonomy

William Buxton

Buxton Design
www.billbuxton.com

(Draft of May 25, 2004)

ABSTRACT

As their name suggests, "projection-vision systems" are systems that utilize a projector, generally as their display, coupled with some form of camera/vision system for input. Projection-vision systems are not new. However, recent technological developments, research into usage, and novel problems emerging from ubiquitous and portable computing have resulted in a growing recognition that they warrant special attention. Collectively, they represent an important, interesting and distinct class of user interface.

The intent of this paper is to present an introduction to projection-vision systems from a human-centric perspective. We develop a number of dimensions according to which they can be characterized. In so doing, we discuss older systems that paved the way, as well as ones that are just emerging.

Our discussion is oriented around issues of usage and user experience. Technology comes to the fore only in terms of its affordances in this regard. Our hope is to help foster a better understanding of these systems, as well as provide a foundation that can assist in making more informed decisions in terms of next steps.

INTRODUCTION

I have a confession to make. At 56 years of age, as much as I hate losing my hair, I hate losing my vision even more. I tell you this to explain why being able to access the web on my smart phone, PDA, or wrist watch provokes nothing more than a yawn from me. Why should I care? I can barely read the hands on my watch, and can't remember the last time that I could read the date on it without my glasses. As a charter member of the baby boom, I suspect that I am not alone.

To make things worse, there is one other thing that seems to be going: my memory. All too often I am confronted by the fact that I can't remember my daughter's phone number, or the address of the company in Vancouver that I am supposed to be visiting.

Here is a true-life experience. Picture a friend and me standing in the street, in the middle of a proverbial Vancouver rainstorm, trying to access the web using his wireless PDA. Our hope is to find the address of the company where we are supposed to be attending a meeting. To make a long story short, yes we go to their web page. No, we could not do anything with it once we got there. To get the picture, just think of me trying to use a display that I can't read to help compensate for a mind that can't remember.

All of this leads to just one of the many reasons that I think that projection vision systems are important: for the first time we have the potential to design portable devices whose display size is not constrained by the size of the device itself. We are on the threshold of being able to create a PDA that I might be able to read, when I want to read it, where I want to read it, with enough information displayed that it is worth reading, and be able to interact with what I am reading without needing the manual dexterity of a neurosurgeon. And best of all, perhaps even be able to do all of the above without my reading glasses!

Figure 1: A Conceptual Representation of a Cell Phone with Laser Projector Display: In addition to the regular LCD display, the cell phone (or PDA) is equipped with a miniature laser projector that can project on any convenient surface. Hence, the display can be much larger than the device, itself, thereby rendering it visible. Furthermore, the same laser that is used for projection also functions as a 3D laser scanner. Hence, users can interact with the information using their fingers on the surface of the projected image. (Image: Symbol Technologies)

How this can be accomplished is illustrated in Figure 1. Besides its regular small LCD display, the cell phone in the figure is equipped with a miniature laser projector. This enables the user to project information from the phone onto any convenient surface, such as a table, elevator wall, or the tray table on the back of an airplane seat. Hence, the projected image can be significantly larger than the cell phone from which it comes. While one has been able to access documents like faxes and web sites from a smart phone for quite a while, there is now the potential for them to be viewed full size, such as A4.

Because it is a laser projector, there are no lenses involved, so the image is always in focus.

Furthermore, the laser projector can be coupled with a camera. What this combination provides is the capability to enable to projector to simultaneously function as a scanner. After all, the projected image is “structured light”. The camera “knows” what is being projected, so it “knows” what it is supposed to look like. So by comparing what it sees with what it is supposed to see, it can do a number of useful things. For example, it can see if your finger is in the image or not, and if so, where, thus enabling the projected image to function much like a touch screen. Thus, “touching” a link on a projected web page would cause that link to be followed. Or, drawing on the projected image would enable one to annotate the displayed document, perhaps even with synchronized voice annotation (since it is a cell phone). Significantly, this is independent of whether the person doing the touching is you or me. Consequently, we see that this most private and personal of devices, is also transformed into a ubiquitous interactive device that supports collaboration.

The vision side of things facilitates other things, as well. For example, it enables the system to automatically do keystoning correction, that is, correcting for distortions caused by projecting onto a surface from an angle. It also has the potential to turn the device into an imager capable of capturing the geometry of the projection surface, along with its true colour (if it is a 3-colour laser projector).

Many of the scenarios put forward by so-called "visionaries" are fantasy at worst, highly speculative at best, and only practical a long way off (if ever). Projection-vision systems are an exception. As we shall see, there is already a strong base of research on which to build. They address some real problems, and many of the enabling technologies are already becfor example, is based on the technology found in conventional bar-code scanners, and therefore has decades of manufacturing and engineering practice behind it.

The scenario that I have given is just one example from a much richer palette. It is just a teaser to motivate you to read further.

SOME CONTEXT

Since the Xerox Star 8010 workstation (Johnson, *et al*, 1989) was first shown in 1981, user interface design has been dominated by the graphical user interface. This is so much the case, that I have come up with what I call *The Rip van Winkle Hypothesis*:

> *If Rip van Winkle had a drivers license, and went to sleep in 1981 after having used a Xerox Star, and woke up today, he would be just as able to "drive" today's computers as drive today's cars.*

Yes, both cars and computers are more powerful, and have wonderful technological advances. But the fundamental model of design and use has not evolved significantly. For the car, this may be a good thing, since it was already a mature technology by 1981. But what about the computer? Certainly the designers of the Star (Johnson, Roberts, Verplank, Smith, Irby, Beard & Mackey, 1989) made a spectacular contribution to the usability and accessibility of computers; however, I am convinced than none of them felt that with the introduction of the Star that the job was done.

That it is not done, and that we have significant problems confronting us, can be seen by considering what has happened in the interim along a number of dimensions, such as the emergence of wireless, mobility, and display technologies. Let us briefly look at the last of these, as an example.

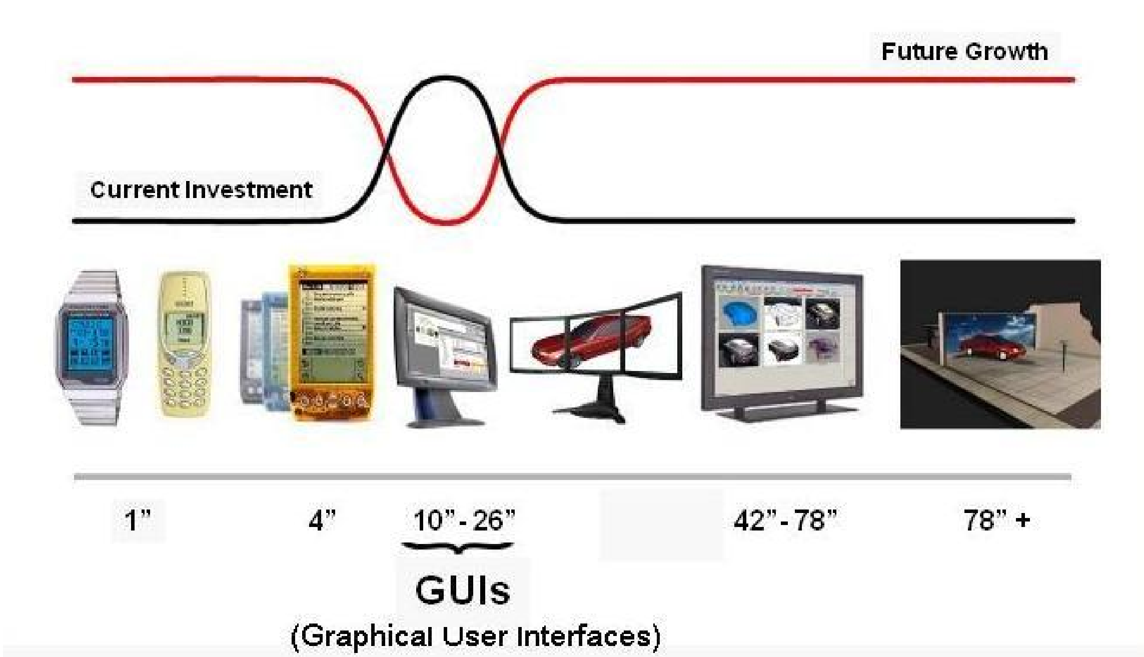

Figure 2: Current expertise and future potential of user interfaces plotted as a function of display size. Note the emergence of both very small and very large displays. Note also that as we move from displays in the 10"-26" range, that the standard GUI becomes ever less useful or relevant.

Figure 2 illustrates a movement towards both smaller and larger displays than those that have traditionally been used on the desktop. The small displays are a key component in the growth of portable electronic devices, such as PDAs, digital cameras, watches and cell phones. They are exploiting mainly LCD technologies, but we are starting to see OLED displays appearing in this sector (see, for example, Kodak's Lifestyle 633 camera). Up to 80", the larger displays are dominated by plasma and LCD technologies, and beyond that, projection technology dominates the market. In automotive design studios, for example, it is not unusual to find high-resolution rear projection screens that are greater than 20' diagonal (Buxton, Fitzmaurice, Balakrishnan. & Kurtenbach, 2000).

What becomes clear in this, very quickly, is that the applicability of the GUI rapidly breaks down as display size goes much below 10" or above 26". Your PDA simply doesn't have enough screen real-estate, or pixels, to support the conventional desktop metaphor. (See, for example, Raghunath & Narayanaswami, 2002). On the other extreme, imagine trying to access the *File* or *Edit* menus off of the top-of-screen toolbar when using a wall-mounted four-foot high electronic whiteboard display. It just won't work.

These new display technologies pose new challenges, as well as opportunities.

- (while can expect erosion of use of projection in larger displays due to emergence) <- why relevant? Why not include projection (LEP) and price performance, as well as e-ink, NED, OLED and LEP, they will still have a role in both directions (add refs, incl Sci American, and Philips …

- unique: display larger than device (espe impt for small)

- can display superimposed over physical objects, i.e., don't need "clean" surface" AR

- coupled with vision, provide unique important, but subtle factor seeing display technologies moving n two trends have emerged insofar as display size is concerned. One is towards smaller displays. Coupled with miniaturization of electronics and wireless, this is helping drive the move towards ever more portable devices, such as PDA's, cell phones, etc. For example, see.

The other trend is towards ever-larger displays. In automotive design studios, 20 foot displays are not uncommon..

through developments in both projection (such as DLP), flat panel Plasma, OLED & LEP) technologies, and e-ink, is towards ever larger, thinner displays at ever lower costs. What is clear is that the further we go down the path towards either large or small displays, the less relevant or applicable is the ubiquitous GUI. This will drive future growth in new approaches to user interface design.

Projection-vision systems are a distinct and important subset of this overall dynamic. Their main contribution is in how they afford an expansion of where and how interaction with electronic devices can take place. They reflect the view that we have stated elsewhere (Buxton, 2001) that it is through innovation in input and output transducer technology that we can most transform the design space of user experience with technology.

In what follows, we will make heavy use of examples. The goal in doing so is to tease out dimensions which are meaningful in distinguishing projection-vision systems, as well as distinguishing the different approaches within the general class. It is hoped that the examples will also help to give a sense of the practical significance of the distinctions identified.

BASIC DIMENSIONS

Addition: Break down and do projection and vision independently first, then combine

Projection:

- larger than device
- remote from device
- superimpose
- can be invisible
- tile
- replaced by oled where static and clear BG
- brightness problems

Vision

- spatial & multiD not fruit fly
- non contact
- remote
- no physical state

Through our analysis of various systems, we arrived at a number of dimensions that appear to be meaningful, in terms of understanding and distinguishing projection-vision systems from the perspective of human usage and experience. Our initial attempt to summarize these can be seen in the following:

- **Projection & Vision**: Does the system support both projection and vision?
- **Idiomatic interaction**: Is the interaction supported particularly idiomatic to the technology, or could it be implemented using another technology such that the change would be essentially transparent to the user?
- **Superimposed & Registered**: Is the field of view of the vision system the same as, and registered with, that field of projection?
- **Richness of the vision system**: does it sense just a point? In 2D or 3D? Or can it capture gestures, documents, or other high degree of freedom date??
- **Target of Interaction**: Is the system for technology mediated human-human interaction, human-Computer interaction or both?
- **Computer**: Does the system do any computation on what is seen or projected, other than purely pragmatic things, such as compression?
- **Camping or parked**: Are the viewed and projected surfaces fixed (parked), or temporary (camped), or in between?
- **Capture Data**: Does the system capture the human user, physical artifacts in the scene, and/or the larger background physical or social ecology of the space?
- **Clean *vs* superimpose projection**: Are the projection and vision systems independent of objects in the field of view, or are physical objects in the projection/vision field significant to the interaction?
- **FG *vs* BG**: Is the system intended to function in the periphery, or background, or are the interactions with it foreground intentional actions? Or does it support both simultaneously, or sequentially?
- **Visible projection**: Is what is projected visible to the human, or is it there in order to support the vision system (as in structured light)?
- **Synchronous or Asynchronous**: Do the vision and projection systems work in series or in parallel?

It is almost certain that this list is not complete and that there are other (perhaps better) ways to categorize things. While acknowledging inevitable omissions and weaknesses, it is hoped that the discussion that follows will render the list useful, even if only to pave the way towards a better analysis by others in the future.

We will now proceed to work through these points one-by-one. Our approach will be by way of examples. These, it is hoped, will help frame our points in both an historical and conceptual context.

KEN KNOWLTON'S SOFT KEYBOARD

An early system that laid part of the conceptual foundation for projection-vision systems was developed at Bell Labs by Ken Knowlton (Knowlton, 1977a; 1977b). This system enabled the functions and labeling of the keys of a physical keyboard to be changed dynamically. This was accomplished by placing a half silvered mirror above the hands, along the optical path from the typist's eyes and the keyboard. Since it was only half silvered, users were able to see their hands on the keyboard. However, they were also able to see labels that were reflected by the mirror from a CRT mounted above.

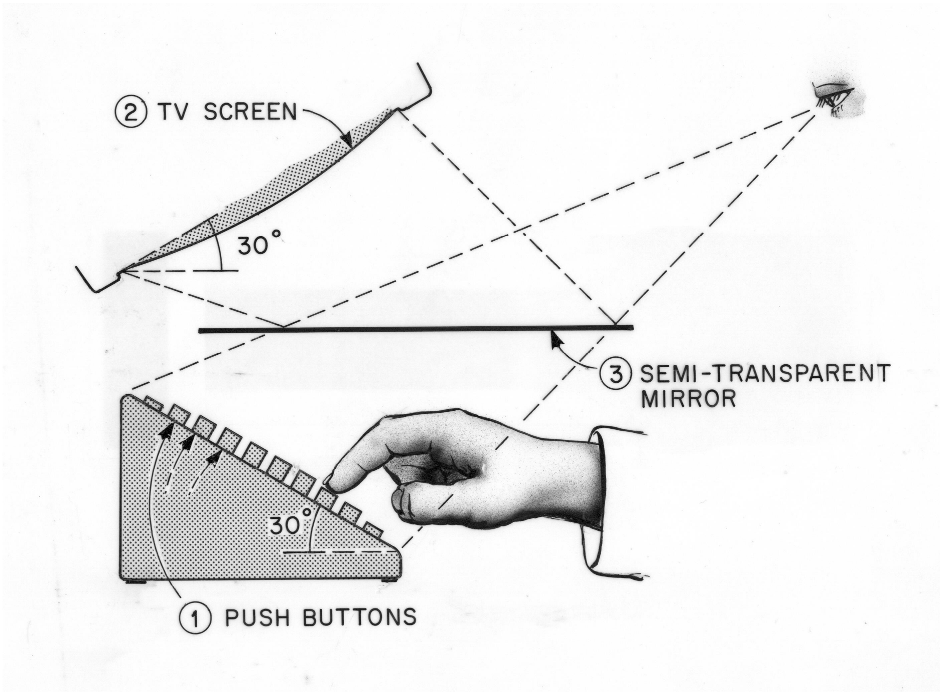

Figure 3: Schematic of Knowlton's keyboard. The image from the CRT was reflected back to the user's eyes by the half-silvered mirror over the keyboard and the user's hands. The effect was that labels for the keys could be changed dynamically, depending on the state of the system, and appear above, rather than below the hands. (Image: Ken Knowlton)

Through this configuration, the labels on the keys were under control of the computer, and could therefore be changed according to the current state of the system.

Figure 4: Key Labels Superimposed Over Fingers. (Image: Ken Knowlton)

While there was no computer vision component to the system, and we are somewhat stretching the notion of projection by including reflection, this system's legitimate ancestry to some of the modern projection-vision systems will become clear in examples that follow.

THE LIGHTPEN

Lightpen were one of the first graphical input devices used in interactive computing. They were first used around 1957 on the TX-0 computer at Lincoln Labs, MIT. Their value in supporting interactive graphics was established by Ivan Sutherland in his pioneering *Sketchpad* system (Sutherland, 1963).

Figure 5: Ivan Sutherland at the console of the TX-2 using his Sketchpad System at Lincoln Labs, MIT, 1963. He is holding a lightpen in his right hand, which he used to interact directly with the screen. The right hand image is a close-up of the lightpen being used for graphical input.

Lightpens consist of a stylus with an optical sensor mounted in the tip. When held against a CRT display, the sensor is used to determine when the display's cathode ray passes by the tip. When the cathode ray is detected, the lightpen determines where the cathode ray was pointing at the time, and uses those coordinates as input to the computer system, thereby performing much the same function as a mouse on a modern day computer.

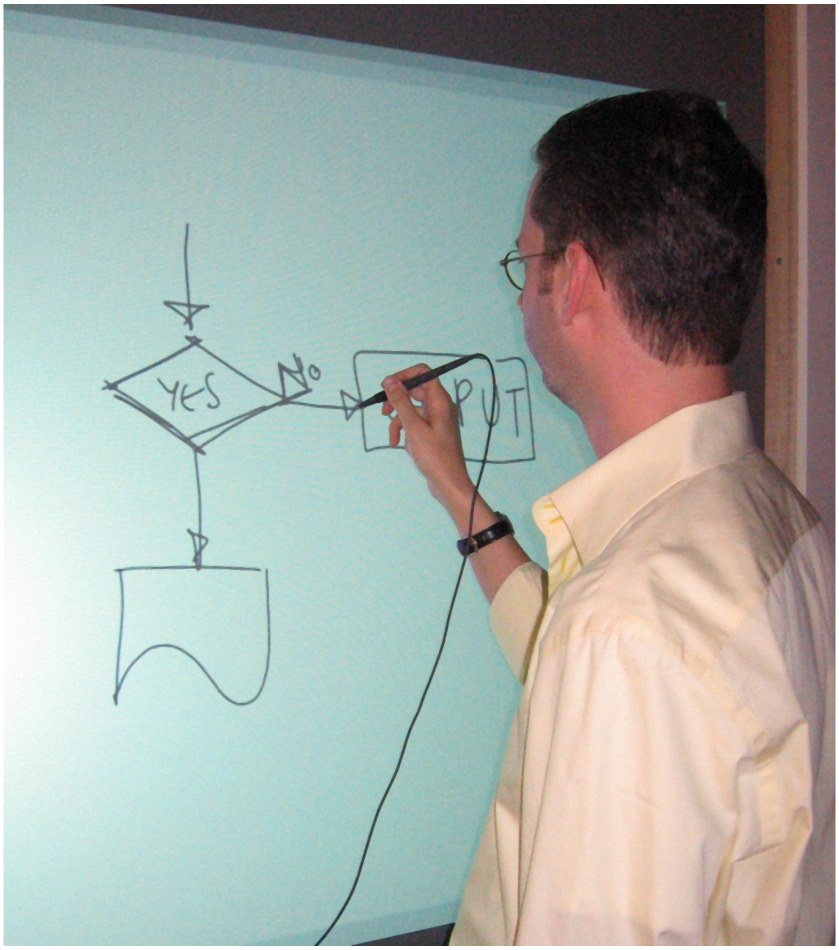

Figure 6: A Lightpen Used With A CRT Projector

While lightpens are generally associated with CRT tubes, they work with CRT (as opposed to LCD or DLP) projectors, since the same fundamental scanning cathode beam is used in both. Few people ever made this leap, which actually made pen-driven electronic whiteboards available years ago, simply by augmenting an existing rear-projection system with a $200 lightpen. (Photo: Azam Khan)

While the optical sensor is clearly a vision system, albeit a very specialized one, the CRT display is not a projection system. However, over the years, a number of labs, including MCC in Austin Texas, and the Ontario Telepresence Project at the University of Toronto figured out that lightpens worked well with rear projections systems, as long as a CRT projector was used. Hence, an easy and inexpensive way was provided to interact with large format electronic whiteboards. The lightpen thus deployed constituted one of the first projection-vision systems.

XEROX PARC LIVEBOARD

In many ways, the lightpen was poorly named. Since the pen sensed, rather than emitted "light", it might better have been called the "eyepen." Be that what it may, a true "lightpen" (in that it emitted light) was developed at Xerox PARC to supporting graphical interaction on a large rear-projection whiteboard, the *Liveboard* (Elrod, et al, 1992).

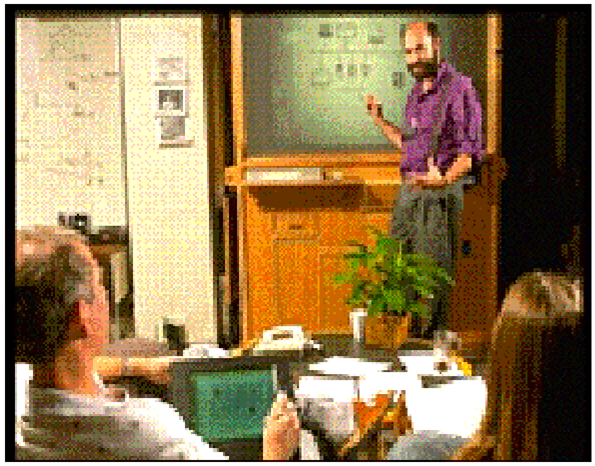

Figure 7: XEROX Liveboard: A large form factor rear projection electronic whiteboard, controlled graphically using a wireless stylus.

The Liveboard used used a wireless stylus that had an LED on the tip. When the pen was held against the front surface of the screen, its presence and location were captured by an optical sensor that was mounted on the rear projector. From the conceptual level, the Liveboard was essentially an improved implementation of the lightpen/CRT rear projection system already discussed, with the added advantage that the stylus was wireless (a significant improvement from the perspective of the user).

PROXIMA CYCLOPS

Following along this same trajectory, we arrive at one of the first commercial systems that was a true projection vision system. It was called *Cyclops*, (Hauck, 1996; Marshall, Hauck, Shapiro, Busch & Stevens, 1996), and was brought to market by Proxima. It was another approach to implementing an electronic whiteboard. However, in this case, a front projection system was used.

Proxima made overhead projectors, and in the early 1990's these were starting to be used in conjunction with LCD panels in order to project computer displays on the wall. This gave rise to the problem of how one might interact with the projected image. In 1993, Proxima came up with the idea of mounting a camera on the head of the projector, and using a vision system to detect and respond to the image of either a red laser pointer, or LED that appeared in the field of view.

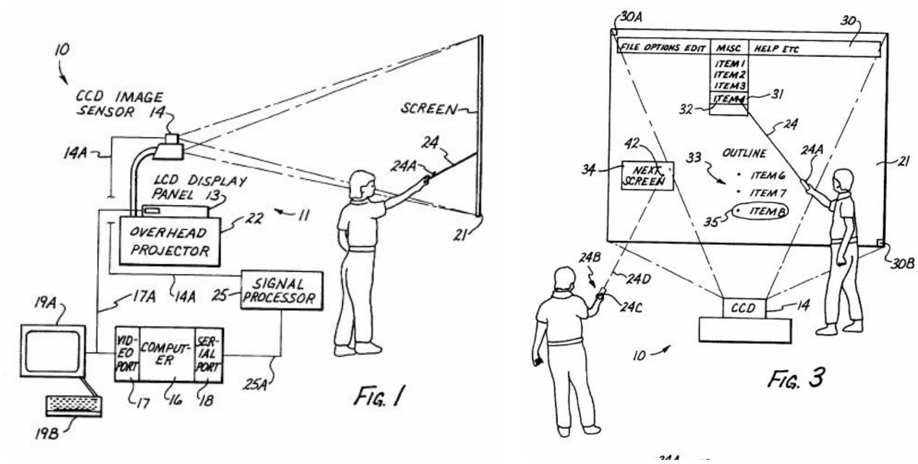

Figure 8: The Proxima *Cyclops* System. The light of an LED mounted on the tip of a pointer, or from a laser pointer, was sensed by a camera mounted on the projector. The camera was calibrated so that its field of view corresponded to the field of the projected image.

The result was something like a light pen, except the optical sensing was on the projector rather than on the stylus, and the property sensed was a particular wavelength of light. (The camera used a band-pass filter tuned to the wavelength of the colour of the pointer light so as to facilitate distinguishing light from the pointer from that emanating from the projector.) Furthermore, two intensities of light were used by the pointer, the brighter of the two being used to indicate a button down state for the pointer, thereby enabling clicking and dragging.

Of the systems discussed thus far, the Cyclops is the first that can be considered as being somewhat mobile, since the overhead projector could (with some effort) be moved from place to place, and was not tied to any particular projection surface or room. Hence, it is our first example of a projection vision system that *camps* rather than *parks* at a particular location.

SMART TECHNOLOGIES

Another, more recent approach to using projection-vision technologies to implement electronic whiteboard functionality is that offered by SMART Technologies. In their case, either front (as illustrated in Figure 9) or rear projection is used for displaying the data. Using their "Digital Vision Touch (DviT) Technology", small cameras (illustrated in Figure 10) are mounted in the corners of the display area so as to view the area directly over the projection surface. Vision software is then used to detect and recognize interactions by the hand, or a stylus, with the projected data.

To the user, in most ways the approach is like a large touch screen. If using a stylus, it is comparable to using a wireless lightpen, or the Xerox Liveboard. What is important about using imaging is that there is no overlay to interfere with the image quality, and unlike an overlay, the technology scales to various surface areas. Finally, the imaging technology has the potential to manage more than one point of contact simultaneously, thereby supporting multiple users, or a single user using two hands. As well, it has the potential to discriminate between narrow points of contact, such as a finger, and wide ones, such as a hand flat on the surface, and use these to establish are broader repertoire of gestures to use in interaction.

Figure 9: Electronic Whiteboard: The camera system shown on the right is used to capture information on and interactions with the whiteboard. The projection system can be used to project data onto whiteboad. (Image: SMART Technologies, Inc.).

As with virtually all systems that rely just on imaging to detect position, as well as event detection, success depends on the reliability and responsiveness of the system in “agreeing with” the user as to if and when it actually touched the surface, or if the finger moved. Problems with the former are a bit like the classic chess player’s argument “You touched that piece!”, “No I didn’t.” As to the latter, especially when standing, a finger can be touching the surface without moving position, and the rest of the finger still have a range of movement, such as if the user shifts body position to get out of the way of the screen, while holding the finger in place. If the devil is in the detail, then here are the devils that determine the effectiveness of these systems for everyday use.

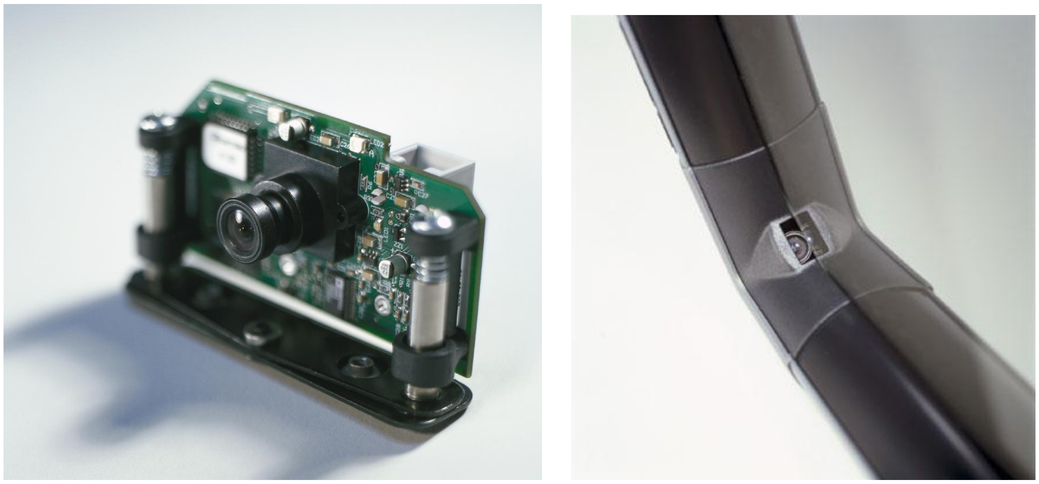

Figure 10: SMART Technology's Digital Vision Touch (DviT) Technology: Digital video cameras (left) are mounted in corners of the projection area (right) and vision software is used to determine the occurrence and location of interaction on the board. The technique avoids any overlays that may affect the image quality, or be impractical, due to the size of the projection area. (Images: SMART Technologies, Inc.)

VIDEO CONFERENCING

Figure 11: Videoconferencing with Projected Image. While cameras are a standard part of videoconferencing, projection is not. Yet, as seen in the photo, rear projection brings life-size scale to the remote participant. This helps bring balance in power to the conversation. It also means that the remote person is defined by their silhouette, rather than by the bezel of the display, which is out of peripheral vision. (Photo: Ontario Telepresence Project, University of Toronto)

Video conferencing, where the remote person is presented by projection, such as illustrated in Figure 11, is another example of projection-vision systems. It is a class where there need be no computation on the signals, except perhaps for compression, where the camera is not pointing at the projection, and where the purpose is technology mediated human-human communication, rather than human-computer interaction (Buxton, 1995).

Projection brings two things to this type of conferencing. The first is an improved sense of presence and balance in the conversation, largely due to the remote person appearing life size. This seems to give their presence "weight" that is not there when using a smaller CRT monitor. Second, if the projection surface is close to the viewer, as in Figure 11, where it is on the other side of the desk, the edges of the screen are near the edge, or out of the viewers peripheral vision. The result is that the remote person becomes defined by the silhouette of their body, as opposed to the bezel of a screen. The impact is powerful in how it helps break the experience away from previous experience of watching television.

KRUEGER'S VIDEOPLACE

With the lightpen, Liveboard and Cyclops, for example, the "vision" component involved capturing a signal that could serve the same purpose as a mouse on a conventional system, i.e, a device with which to point, click, select and drag. In the 1970's, Myron Krueger developed a projection-vision system which accomplished this, and more, by looking at the human body, rather than a physical tool or some other intermediary (Krueger, 1983,1991; Krueger, Gionfriddo & Hinrichsen, 1985).

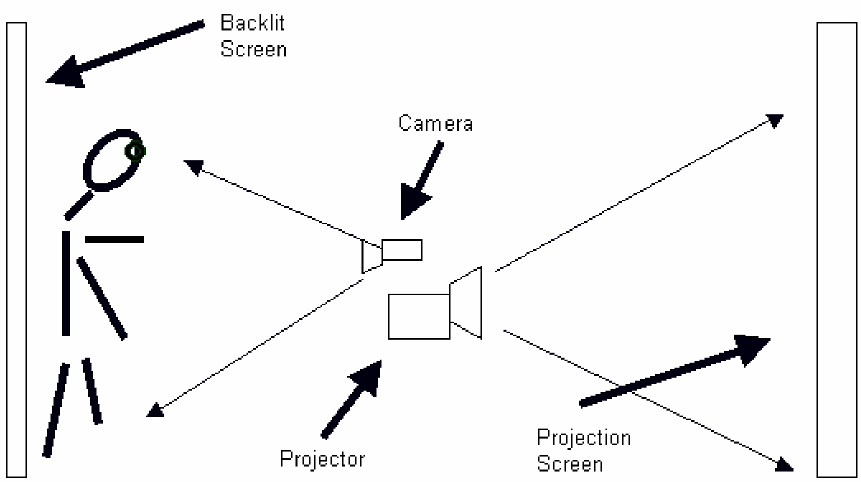

Figure 12: Videoplace by Myron Krueger. The camera captured the body's silhouette, contrasted against a back-lit screen, and captured features from it, such as finger positions and gestures, and used these to control the software projected on a screen in front of the user, generally with a digital representation of the silhouette superimposed on the computer generated graphics.

In his system, depicted in Figure 12, the user stands in front of a back-lit screen. A video camera is pointed at the user and fed into the computer. Because of the high contrast between the image of the user and the background, the computer is able to quickly make a silhouette of the user. Working from this, the computer looks for specific features, such as the position and pose of the index finger, and uses this information to control interaction (Krueger 1983, 1991; Krueger, Gionfriddo & Hinrichsen, 1983).

Depending on what program is running, a copy of the silhouette is projected on the screen in front of the user, along with the appropriate output of the program resulting from the user's input. Figure 13 illustrates two applications. The one on the left shows the user finger painting, and the one on the right "typing", or entering text. In both cases, the user interacts by pointing with the index finger.

Figure 13: Drawing and Typing Using Videoplace: In the left image, the program lets you "finger paint." You apply digital "ink" using your extended index finger. Likewise, in the right hand typing example, you enter a character by pointing at it with your index finger.

To this point, all of the systems discussed involved capturing a signal that was analogous to the information provided by a mouse. That is, they provided an (x,y) coordinate pair giving position as well as one bit of state, to indicate if the "mouse button" was up or down.

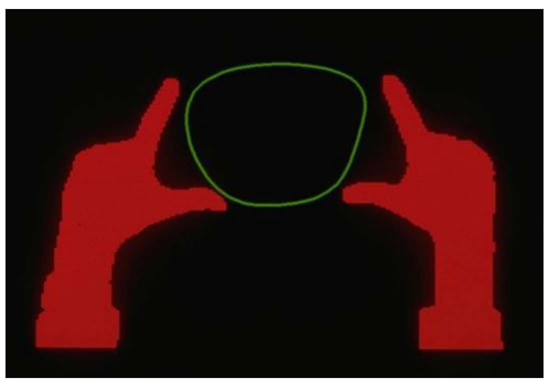

Figure 14: Rich Manual Gestures in Videoplace**:** Here, the tips of the two index fingers and the two thumbs are being used to control the shape of a circular type object. The user could dynamically squeeze or stretch the edges in order to modify the form.

Because it captured the user's silhouette, rather than just the position of a stylus or finger, Krueger's system had a much broader repertoire of things that it could react to. Essentially, virtually any meaningful gesture that could be articulated by such a silhouette, using hands, body, feet, or all of the above, was legitimate fodder for his system. This is illustrated in Figure 14, where two fingers and two thumbs are being used simultaneously to control the shape and position of a circular type figure. The potential of using vision to enable users to use a rich, natural set of gestures, unencumbered by gloves or complex devices, was firmly established by Krueger. Unfortunately, despite being publicly known for 20 years, this work has not had much impact on interaction in general. Recent dramatic improvements in the price/performance characteristics of video, computational and display technologies mean that it may be time to revisit this work with new eyes.

TANG & MINNEMAN VIDEOWHITEBOARD

Another system that is similar in concept, but completely different in implementation and function to that of Krueger, is the *Videowhiteboard* developed by John Tang and Scott Minneman at Xerox PARC (Tang & Minneman, 1991a, 1991b). This was a collaborative drawing system, where two people in different locations could both draw on the same virtual whiteboard.

The system was stunning for its simplicity. In fact, there was no computer involved in its implementation. There were just two video cameras and projectors, set up as illustrated in Figure 15. Each user drew on their "whiteboard" (actually a rigid rear-projection screen). A camera was pointed at the back of the screen and was able to capture, due to the contrast with the back lighting, both what was written, as well as the shadow of the person. This was projected onto the back of the other person's whiteboard, so that they could see both what was written/drawn, as well as the shadows. They could then draw on their board, integrating their work with that from the other person. And, a camera would capture their work and project it onto the back of the other person's screen. Thus, each could work on the "same" display, but only erase lines which they had made themselves.

An example of what this looked like is shown in Figure 16. Note both the similarities and differences to the examples of Krueger's work.

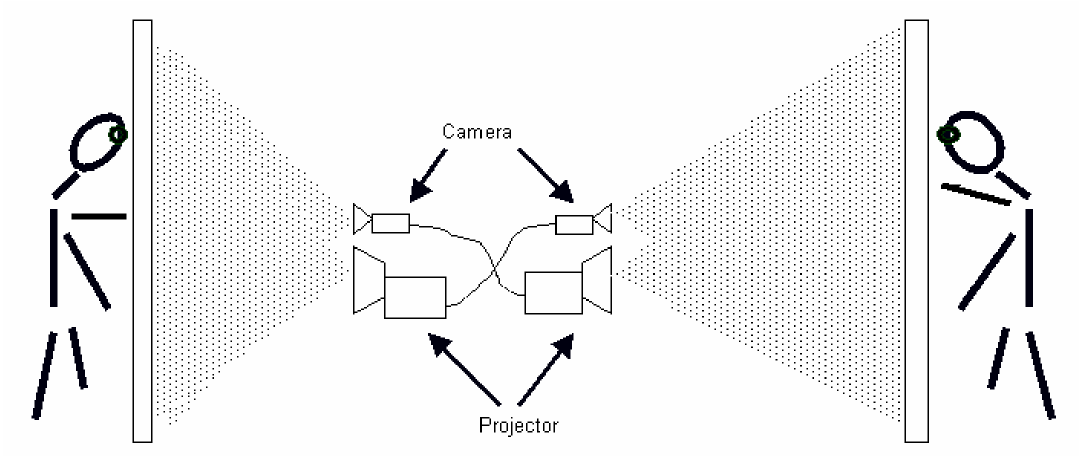

Figure 15: Videowhiteboard Schematic: Users draw on their side of the whiteboard (actually a rigid rear projection screen) using standard whiteboard markers. A camera is pointed at the back of the whiteboard, and its output projected onto the back of the other, and vice-versa. Hence, both the marks and shadow of each person were transmitted to the other. Polarizing filters were used to avoid video feedback.

While Krueger also did remote drawing systems, and had each person represented by a silhouette, Videowhiteboard had some significant differences. First, it was not a gesture recognition system. The collaborators drew using conventional whiteboard markers. Second, the shadows were exclusively for human-human communication. They had no impact on the system technically. (Remember, there was no computer.) All that they did was give a sense of presence to the remote person, and this was very much enhanced by the shadows being graduated, getting darker and sharper as the person approached their board to write or gesture to the remote person. The resulting sense of depth and distance from the work was perceptually very different from Krueger's system.

Despite the absence of a computer, Videowhiteboard certainly is a projection-vision system. It managed to work without a computer since its sole purpose was for technology mediated human-

human communication. Interacting with the technology was not a relevant issue. It also is the first system that we have seen where there is a clear distinction between two levels of communication: intentional marking actions, and referential gestures that related primarily to what was drawn. In some ways this was part of Krueger's work. However, since he used hand gestures for communicating with both the remote user and the machine, the potential for confusion existed. On the other hand, by using markers for drawing, rather than hand gestures, these two classes of interaction were clearly distinguished in Videowhiteboard. The cost, however, was being restricted again to marking with a point, rather than the high degree of freedom gestures that Krueger supported.

There is no right or wrong here. What is most appropriate obviously depends on the context. However, there are two more points worth noting.

First, for me, Videowhiteboard is as interesting for its methodology as it is for what it did. What economy of design and effort! It is rare to see researchers get so much insight and experience from so little implementation effort. They spent their time and creativity in thinking about what to do, rather than rushing into building something. Their work is the interaction equivalent of a sketch: fast, inexpensive, and rich. Reimplementing this system should perhaps be a standard exercise for any student of interaction design.

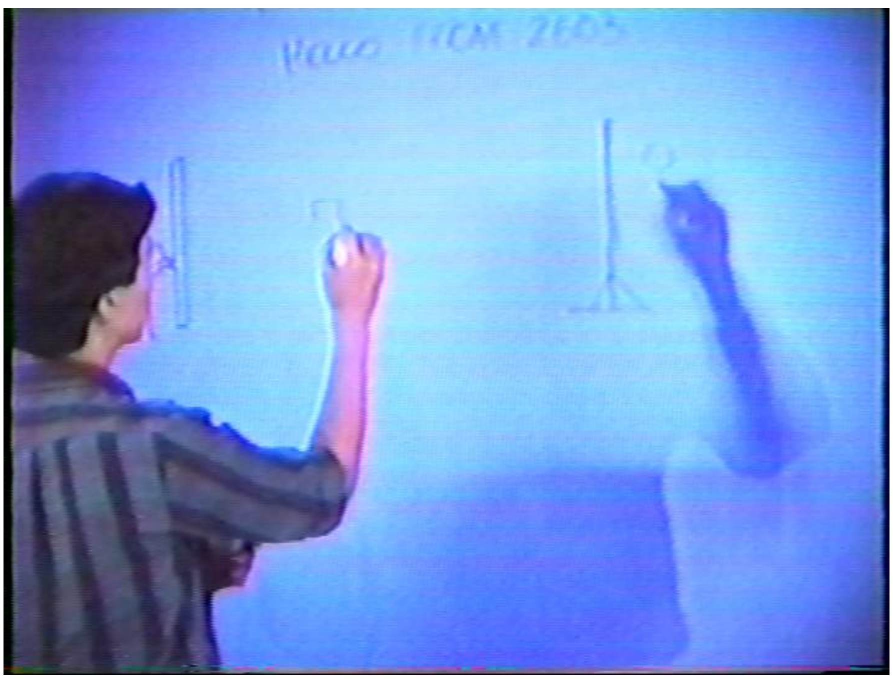

Figure 16: Videowhiteboard: The shadow of the remote person is shown. They appear to be standing just on the other side of the screen. Note how the intensity and sharpness of the shadow indicates how close they are to the board.

Second, there is nothing in the Videowhiteboard that could not be brought over to the digital domain. That this is technically feasible is clear simply by looking at Krueger's work. That there has been little or no follow-up to this work is, to me, stunning.

PIERRE WELLNER: DIGITALDESK

The DigitalDesk (Wellner, 1991, 1993) is an example of using a projection vision system to create what has become known as an *Augmented Reality* system. That is, a system where the normal physical objects and artifacts are augmented through computational means. In this case, on the one hand, virtual objects appear on the user's work surface,

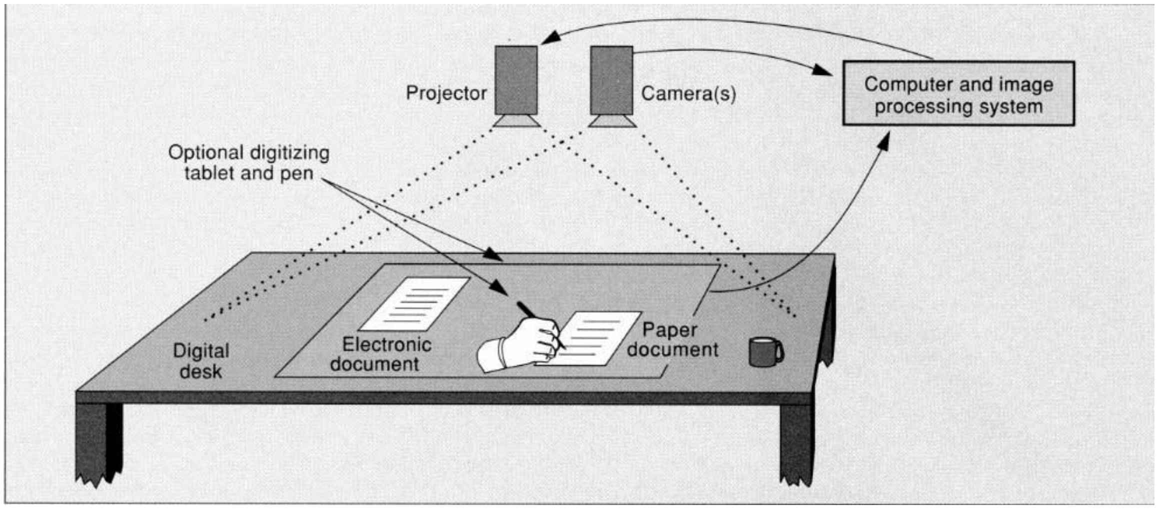

Figure 17: Wellner's DigitalDesk. This system uses a projection vision system to implement an augmented reality. The system projects over the workspace, including documents, as opposed to a blank screen. The purpose is to support interaction on a common surface with both physical and virtual documents and devices. Input by the user is accomplished using signals captured by the video camera, which is also pointed at the desk. (Image: Pierre Wellner)

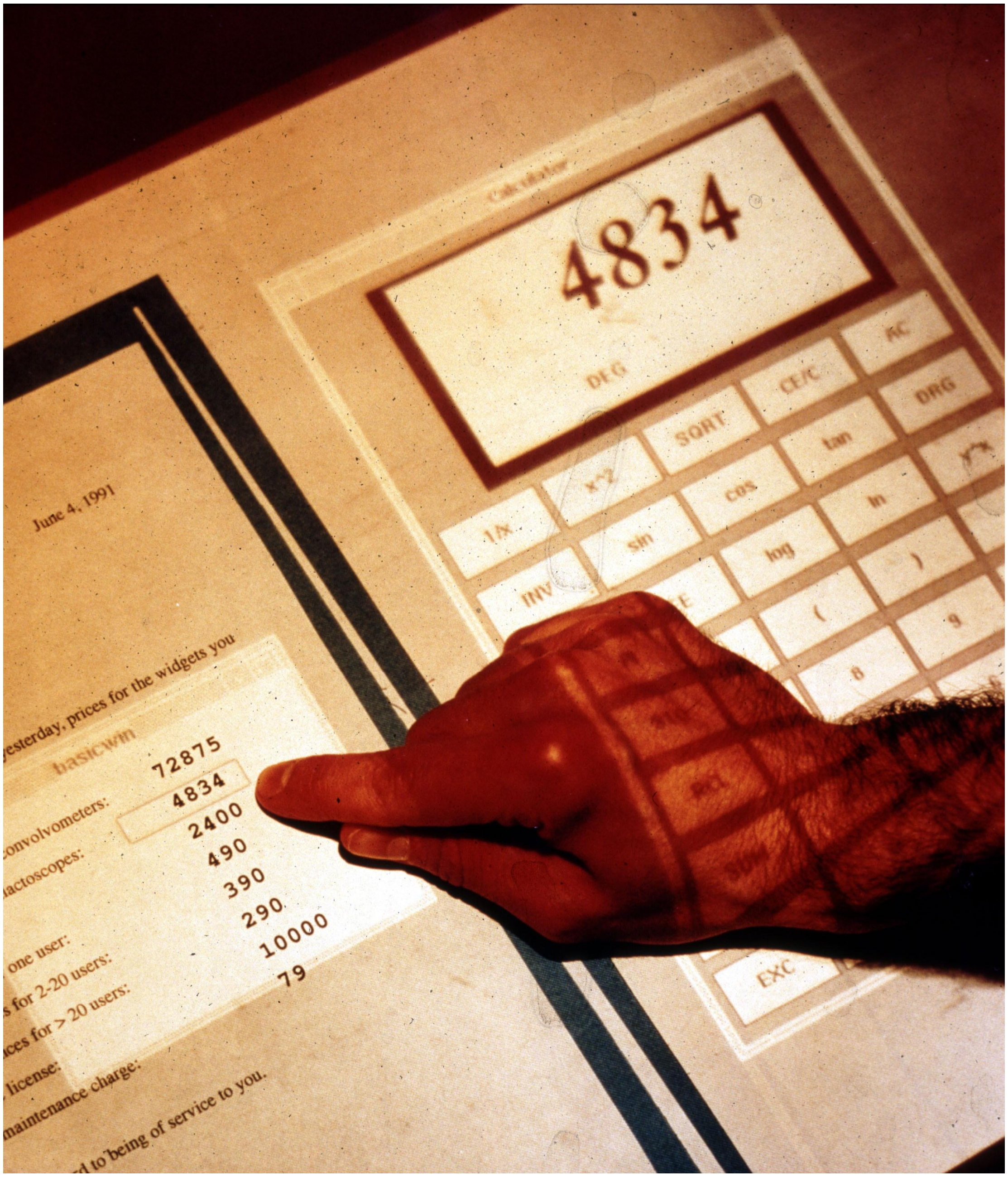

Figure 18: Bridging the physical and the virtual. **The user's finger is tracked by a camera and "sees" that it is pointing to a number on a physical piece of paper, in this case, 4834. The number is "read" (OCR'd) using the camera, and entered by the computer into the virtual calculator projected on the desk. The user can then perform calculations on this number by tapping on the virtual keys of the calculator with the finger. Notice the similar genealogy to the work of Knowlton, discussed earlier. (Photo: Pierre Wellner)**

iLight Fuji Xerox

Writing on remote whiteboard. Capture whiteboard with camera, correct keystoning, annotate with tablet pc, send back, and project over physical board. Way to remotely annotate. Jonathan Foote, FX Palo Alto Laboratory.

Virtek Vision International Inc

Virtek Vision is an example of how projection/vision systems can scale. Most of the technologies discussed on this page are concerned with providing relatively large (on the order of 10-30 c.m.) virtual display and input transducers to small portable devices. Virtek works at the opposite extreme. They use laser projection systems for very large scale functions, such as in manufacturing large wooden or metal parts.

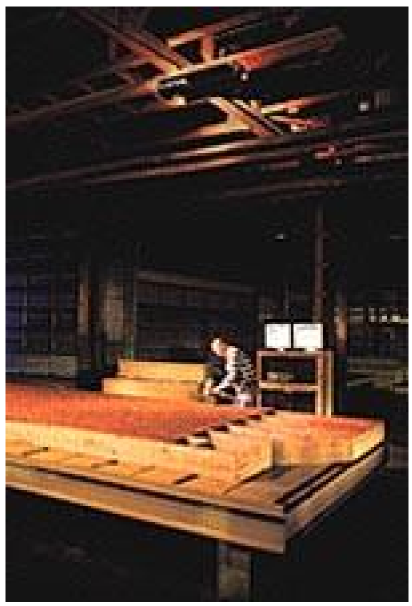

Figure 19: Projection System Used for Templating Purposes. **Here a laser system is used to project the pattern onto large piecework. Later in the process, parts can be inspected using a different laser system. (Image: Virtek Vision International, Inc.)**

In this case, their technology is used to project patterns or templates onto the materials. They also provide technologies for using laser-based vision systems to do quality assurance inspection. They are not currently applying their technology to what we would normally call "human computer interaction", and they would probably not consider themselves as enabling "virtual interactive devices," nevertheless, their technology is relevant to the class of system being described.

3DV SYSTEMS Z-CAM

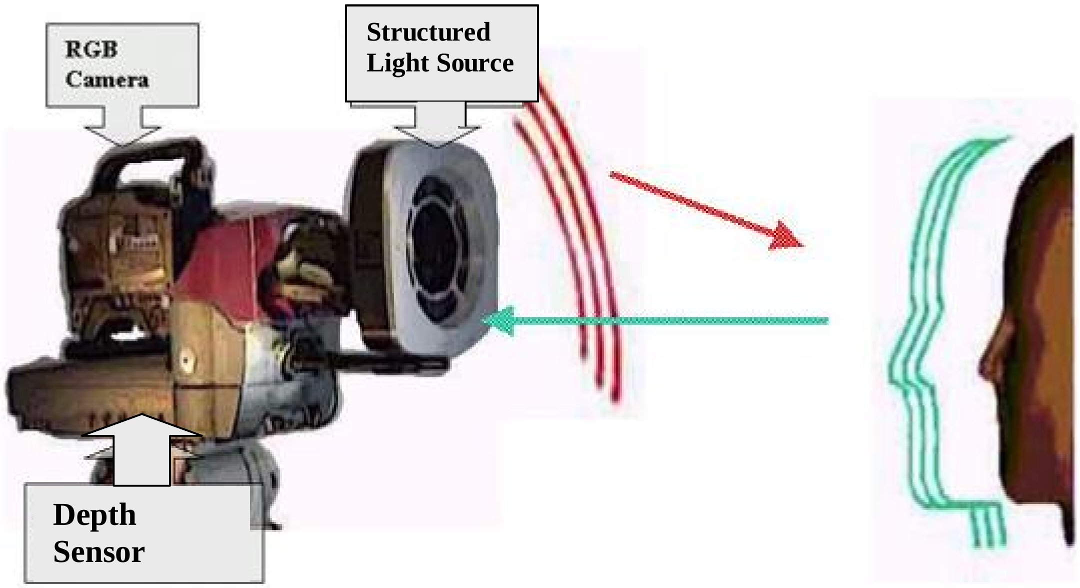

Figure 20: ZCam Object Camera: **A conventional video camera is augmented by a depth sensor and a structured light projector. A known pattern of light is projected onto the subject. It is invisible to the human eye, but visible to the special camera of the depth sensor. The 3D form of the object can then be inferred by the differences between the pattern projected and the pattern seen by the sensor. This information can be coupled with the image captured by the regular video camera thereby enabling the combined system to know the distance of each pixel in the video image from the camera. (Image: 3DV Systems)**

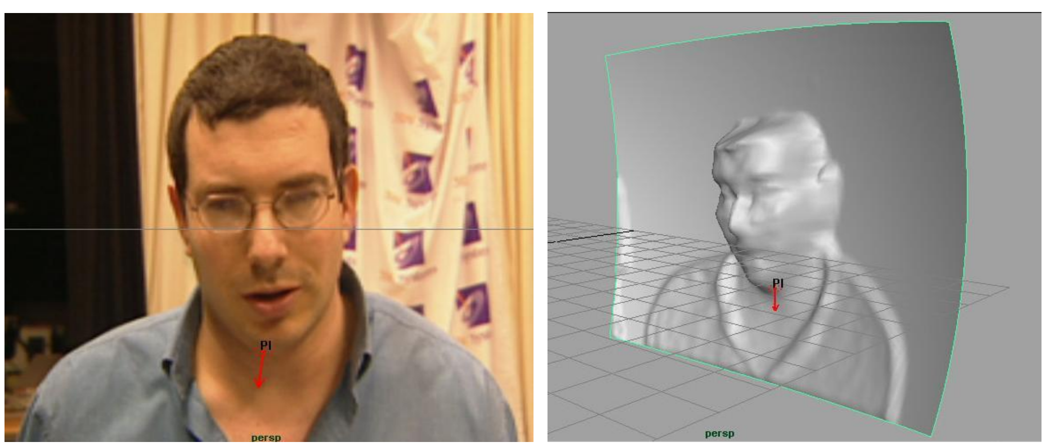

Figure 21: ZCam Output: **The image on the right is the depth information captured by the depth sensor from the image on the left, captured by the video camera. The right hand image is seen from the side in order to highlight the visibility of the depth information. (Image: 3dv systems)**

I/O Bulb

Underkoffer, Ullmer & Ishii, 1999)

VKB

Figure 22: Another Virtual Keyboard: **This unit was demonstrated by Siemens Procurement Logistics Services at CeBIT 2003 in Hanover, Germany. The technology was developed and manufactured by VKB Inc, in Israel. (Image: VKB, Inc.)**

CANESTA

Figure 23: Canesta Projection Keyboard: **The keyboard is projected onto the work surface by a projector mounted at the top of the PDA. The sensor senses the fingers breaking the invisible light fanning out from the light source, close to, and parallel to, the desk top. Using the information from the sensor, software determines that the fingers have broken the fan of light, and where, thereby determining which virtual key has been struck. (Photos: Canesta, Inc.)**

IBM STEERABLE CAMERA_PROJECTION SYSTEM

Pingali, et al (2003)

Kjeldsen et al (2003)

Chou et al (2001) Bluespace

Figure 24: Steerable Camera and Projector. **(Image: IBM)**

Figure 25: Bluespace: Reconfiguring Display Surface in Office Context. **The same steerable projector-camera pair can direct the output display onto any number of different surfaces, including the desk, wall, and meeting table (as well as the floor, etc.). The user(s) can interact with the data displayed in any of the locations, by means of the camera. (Images: IBM)**

Figure 26: Images Without Boundaries: **The frog avitar is projected onto the wall. The camera detects when the hand approaches, causing the frog to jump away. While a seemingly trivial example, perhaps, it is significant not only for the richness of the gesture, but where it takes place, and how remote it is from both a visual and interaction sense, from our experience with conventional video. (Image: IBM)**

SYMBOL

- include comparison with augmented reality with HMD's
- ability to manually do fine registration w/o glasses
- easier to deploy (common hardware)
-

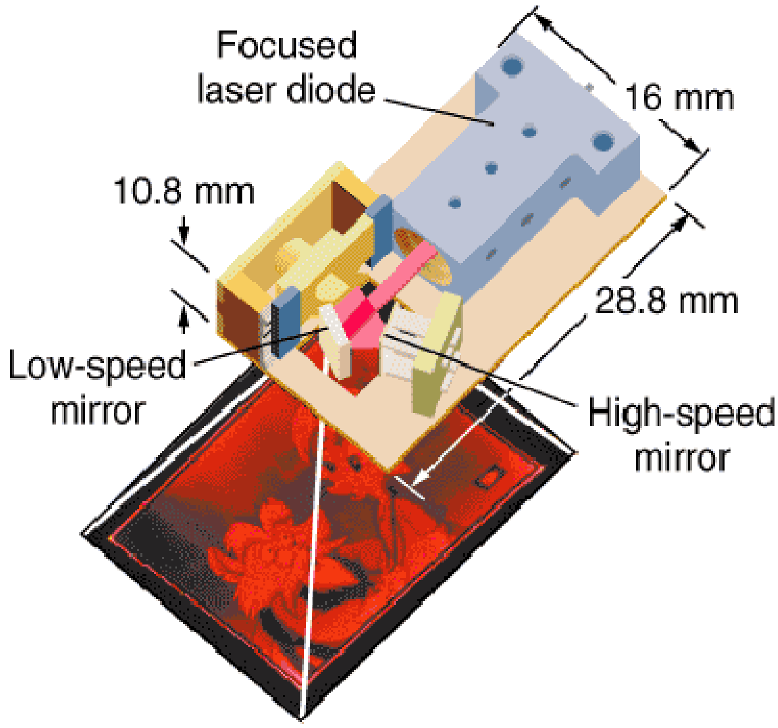

Figure 27: Schematic of Prototype Miniature Laser Projector. **The prototype projector uses a laser to project a video resolution (640x480) monochrome image with 16 levels of intensity (grey-scale). In production, the projector can be expected to be 30%-50% smaller, and would to benefit from economies of scale, use the same kind of laser used in DVD and CD players. (Image: Symbol Technologies).**

Figure 28: Example Output from Prototype Miniature Laser Projector. **Monochrome images can be projected on arbitrary surfaces. Because the projector uses a laser, the images will always be in focus.**

CONCLUSIONS

- Jean Piaget: An Ecological Approach
- Buxton Lemma
- Kranzberg's Laws

ACKNOWLEDGEMENTS

Pierre Wellner of the IDIAPResearch Institute, Switzerland,

Mike Flynn" mike.flynn@idiap.ch

Alistair Hamilton

"Jennifer Lo" <JenniferLo@smarttech.com>

REFERENCS/BIBLIOGRAPHY

3DV Systems: http://www.3dvsystems.com/products/zcam.html

Buxton, W. (1995). Integrating the Periphery and Context: A New Model of Telematics *Proceedings of Graphics Interface '95*, 239-246.

Buxton, W. (2001). Less is More (More or Less). In P. Denning (Ed.), *The Invisible Future: The seamless integration of t*echnology in everyday life. New York: McGraw Hill, 145 – 179.

Buxton, W., Fitzmaurice, G. Balakrishnan, R. & Kurtenbach, G. (2000). Large Displays in Automotive Design. *IEEE Computer Graphics and Applications*, 20(4), 68-75.

Canesta, Inc.: http://www.canesta.com

Chou, P., Gruteser, M., Lai, J., Levas, T., McFaddin, S., Pinhanez, C., Viveros, M., Wong, D. & Yoshihama, S. (2001). BlueSpace: Creating a Personalized and Context-Aware Workspace. *IBM Research Report RC22281 (W0112-044)*.

Elrod, S., Bruce, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., Tang, J. & Welch, B. (1992). Liveboard: a large interactive display supporting group meetings, presentations and remote collaborations. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'92)*, 599-607.

Hauk, Lane (1996). Computer input system and method of using same. US Patent 5,515,079.

IBM (2003): http://www.research.ibm.com/ed/

Johnson, J., Roberts, T.L., Verplank, W., Smith, D.C., Irby, C.H., Beard, M., Mackey, K. (1989). The Xerox Star: A Retrospective, *IEEE Computer*, 22 (9), pp 11 – 26.

Kjeldsen, Rick , Levas, Anthony , & Pinhanez, Claudio (2003). Dynamically Reconfigurable Vision-Based User Interfaces. *Proceedings of the 3rd International Conference on Vision Systems (ICVS'03)*. Graz, Austria.

Knowlton, K. (1975). Virtual Pushbuttons as a Means of Person-Machine Interaction. *Proceedings of the IEEE Conference on Computer Graphics, Pattern Recognition, and Data Structures*. 350-351.

Knowlton, K. (1977a). Computer Displays Optically Superimposed on Input Devices. *The Bell System Technical Journal*, 56 (3), 367-383.

Knowlton, K. (1977b). *Prototype for a Flexible Telephone Operator's Console Using Computer Graphics*. 16 mm film, Bell Labs, Murray Hill, NJ.

Koike, Hideki & Kobayashi, Yoshinori (2001). Integrating paper and digital information on EnhancedDesk: a method for realtime finger tracking on an augmented desk system. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 8(4), 307 - 322

Krueger, Myron, W. (1983). *Artificial Reality*. Reading, MA: Addison-Wesley.

Krueger, Myron, W. (1991). *Artificial Reality II*. Reading, MA: Addison-Wesley.

Krueger, Myron, W., Gionfriddo, Thomas., & Hinrichsen, Katrin (1985). VIDEOPLACE - An Artificial Reality, *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'85),* 35 - 40.

Marshall, R., Hauck, L., Shapiro, L., Busch, J. & Stevens, E. (1996). Method and apparatus for calibrating an optical computer input system. US Patent 5,489,923**.**

Pingali, G., Pinhanez, C., Levas, A., Kjeldsen, R., Podlaseck, M., Chen, H., Sukaviriya, N. (2003).Steerable Interfaces for Pervasive Computing Spaces In *IEEE International Conference on Pervasive Computing and Communications - PerCom'03*. Dallas-Fort Worth, Texas.

Raghunath, M. T. & Narayanaswami, C. (2002). User Interfaces for Applications on a Wrist Watch. *Personal and Ubiquitous Computing***,** 6(1), 17-30.

SMART Technologies, Inc.: http://www.smarttech.com/dvit/

Sutherland, I. (1963). SKETCHPAD: A Man-Machine Graphical Communication System, *Spring Joint Computer Conference*, 329.

Symbol Technologies: http://www.symbol.com/products/oem/lpd.html

Tang, J.C. & Minneman, S.L. (1991a). Videowhiteboard: video shadows to support remote collaboration. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'91)*, 315-322.

Tang, J.C. & Minneman, S.L. (1991b). VideoDraw: A video interface for collaborative drawing. *ACM Transactions on Information Systems*, 9(3), 170-184.

Tomasi, C., Rafii, A., & Torunoglu, I. (2003). Full-size projection keyboard for handheld devices. *Communications of the ACM (CACM)*, 46(7), 70-75.

Underkoffler, J., Ullmer, B., & Ishii H. (1999). Emancipated Pixels: Real-World Graphics in the Luminous Room. In *Computer Graphics Proceedings (SIGGRAPH 99)*, ACM Press, 385-392.

Virtek Vision International Inc.: http://www.virtekvision.com

Virtual Devices Inc. (VDI): http://www.virtualdevices.net/

VKB, Inc.: http://www.vkb.co.il/

Wellner, Pierre (1991). The DigitalDesk Calculator: Tactile manipulation on a desktop display. *Proceedings of the Fourth ACM SIGGRAPH Symposium on User Interface Technology (UIST'91)*, 27-33.

Wellner, Pierre (1993). Interacting with paper on the DigitalDesk. *Communications of the ACM (CACM)*, 36(7), 87-96.

Wilson, Andrew (2004). TouchLight: An Imaging Touch Screen and Display for Gesture-Based Interaction. *Proceedings of the International Conference on Multimodal Interfaces*, 69-76.

NotesFujix P-40U Handy Projector with Stereo Sound.
LL98992
No. 2010131
DC 6V 21W
Factory Code: KKC