Aspect with Invisible Computing

Assignment 2

PhD Proposal

Direct Interaction with Large-Scale Display Systems at a distance

The University of Sydney

School of Information Technologies

Sydney, Australia

Contents

Chapter 1

Aim and Significance of the Project

Large scale display systems spanning an entire wall are widely used in many modern information technology facilities, especially for non-interactive purposes such as presentations. Where they are used interactively, the user interaction devices typically consist of a standard keyboard and mouse. However, there are a number of reasons why these devices are less than optimal for large displays. From the outset, in 1968, Douglas Engelbart developed the mouse to provide a way for users to interact with personal computers [1]. It was never designed to be used in a large display environment. As a result, the mouse only performs moderately well when scaled to large screens.

Imagine a user giving a talk to an audience while their presentation slides are being shown on a large screen behind them. To use the mouse, the user needs to place it on a desk or a flat surface. This constrains the user to stay within arm's reach of the table, thus reducing their mobility. To interact with the system, the user moves the mouse in a horizontal plane across the table, while the cursor moves in a vertical plane of the display. There is a need for users to adjust for small movements of the mouse with a corresponding larger movement on the screen. Thus, time is required for the user to think what action is needed in order for the cursor to reach the desired position on the display. Their cognitive load is therefore increased, distracting the user away from their task. Yet another problem is the need to turn around to see where the pointer is on the large display. Thus the mouse is not optimal for interacting with large displays.

A better approach is a system that allows for direct interaction between the user and the objects seen on the large display, for example, manual manoeuvres to point at the display or to rotate objects by twisting, pushing or turning of the hands. In general, we need a device that allows the user to interact directly with the display without the need for an intermediary device. Such systems are more natural and easier to use.

The aim of the proposed PhD project is to explore ways of interacting with large displays at a distance using a more direct approach. The method chosen is a small handheld pointing device, the size of a TV remote control. A built-in radio transmitter is fitted on both ends of the device and their coordinates in 3D space are transmitted back to a stationary receiver using radio link. The resultant position being pointed at on the screen can then be calculated. In addition, a visible red laser pointer can be fitted on top of this device so that users can see where they are pointing to. The aim of such device is to remove the need for cable, and to remove the need for the line-of-sight requirement in a system which is required to track the position of the pointing device or the need to track the position of the cursor onscreen. The project will also investigate how selection can be done with such device in addition to the two techniques examined in an honours thesis, namely hotspots and gestures[2]. It would also be possible to use such device to act as a model for the object onscreen, so that users can rotate the device in their hand and the corresponding action can be produced with the object onscreen. (Details of these techniques are discussed in Chapter 3). The usefulness of these techniques will be evaluated in a presentation environment.

Apart from presentation environments, it is anticipated that these techniques can be applied to other domains as well.

Graph movement

By using the pointing device to select individual nodes or cluster of nodes, graphs can be directly manipulated in 2D or 3D space using such actions as rotation and transformation. This would provide a more direct approach to relational information visualisation.

Focus and context

Our pointing device can be used to point directly at an area of focus. Particular techniques of potential include the Document Lens [3], Fisheye View [4] and the Hyperbolic Browser [5]. These only require one focal point and can easily be integrated with our interactive paradigm.

Selective dynamic manipulation

Gestures can be used to select arbitrary object sets [6]. These object sets can be manipulated and their appearance changed so that different tasks can be performed.

Collaboration

Multiple users can interact simultaneously using multiple pointing devices [7]. It may be possible to distinguish the different devices by using different radio frequency.

See-through tools

By using two pointing devices and gestures, users can use the Toolglass paradigm to select application tools interactively [8]. By using a transparent tool palette that can be placed between the application and the cursor, minimal cursor movement can be achieved with reduced errors.

Flow menus

Direct manipulation, menu selection and text entry are merged in flow menus so that they can be performed without the use of a button click. This is extremely useful when used with our pointing device [9]. The start and end point of a gesture do not have to be known and it allows consecutive menu selection.

It is anticipated that the outcome of the project will provide the general public with a more direct way of interacting with large displays in the above domains whether they be consumers, corporate users or researchers. It will allow cheaper setup cost because tracking cameras will not be needed and users can move around freely without worrying about occluding the device from the tracking camera. Especially for the research community, it would be possible to investigate optimising techniques such as reducing the latency between the transmitter, the receiver and the actual movement of the onscreen cursor, improving the accuracy of such device, extending methods for button-less selection or 3D manipulation of objects using such devices.

Chapter 2

Related Work

Manipulation is the adjustment or changes made to an object in some shape or form. In terms of information technology, it goes as far back as the 1960s where Seymour Papert at the Massachusetts Institute of Technology developed the LOGO computer programming language. The concept of symbolic computation [10] was adopted to give a visual representation to words and ideas. A graphical representation, a turtle, is used to allow children to type simple commands and observe for the effects of the turtle on screen. This is an example of *indirect manipulation* where instructions are entered into the computer using the keyboard in order to manipulate the turtle. Alternatively, a more direct approach is to control the turtle by interacting with its on screen representation. This is the principle behind *direct manipulation*, a term coined by Ben Shneiderman in 1983 [11]. MacDraw uses this approach to allow users to manipulate lines and shapes on a drawing [12,13] by providing users with a tool palette along the left side of the drawing. Users can then select different tools using their mouse to manipulate the drawing as desired. The manipulation is achieved by using the mouse to interact with the system.

Interaction, as distinct from manipulation, refers to the method which users provide input to the system, and the system provides feedback to the user. The interaction technique used in this project is the pointing device. The traditional keyboard is sufficient for inputting characters or numbers, such as that used by children to enter simple commands in LOGO. However, convenience and efficiency can be achieved if one can point or aim directly at a location we are interested in. Such a system would be more direct and easier to use.

Over the years, various novel input devices have been developed to address the problem of indirectness, with varying degrees of success – a comprehensive overview is outlined in [14]. We now consider some devices that provide a somewhat more direct approach than the mouse.

The *lightpen* is one of the first pointing devices produced [15]. It works by pressing the pen against a CRT display which is operated by a switch on the pen and allows it to capture light produced from the screen. However it is not suitable for prolonged use due to its poor ergonomics, thus it was eventually replaced by the mouse. As discussed in Chapter 1, the mouse provides an indirect method to interact with the computer.

Figure 2.1: Light pen

Light gun technology has been used extensively, especially in the computer gaming industry. It also provides a direct approach to interacting with CRTs and unlike light pens it can operate at a distance. Although the accuracy is maintained, its major drawback is that it must be used with a CRT display. Thus, large-scale displays cannot be used since images are provided by a data projector. Another disadvantage is the requirement of a cable to connect the light gun to the main system.

Figure 2.2: Light gun

The *Polhemus FasTrak* and *Logitech 3D Mouse* belong to a category of industrial strength tracking systems primarily designed for 3D motion tracking, as found in applications such as *CavePainting* [16].

FasTrak is an electromagnetic tracking system that computes the position and orientation of a tiny receiver as it moves through space. The major problem however is its vulnerability to electromagnetic interference and radiation particularly from the monitor. In addition, this system has a limited range of 3 metres and a latency of 4 milliseconds. It is primarily designed for 3D motion capturing in a Virtual Reality environment. The 3D Mouse is another similar tracking system. It uses a stationary triangular transmitter which emits ultrasonic signals to track the movement of a smaller triangular receiver. This resolves the problem of interference from radiation but introduces interference by other equipments that use ultrasonic signals. The system also has a limited range of 2 metres and a high latency of 30 milliseconds. These are typically used for CAD object manipulation and Virtual Reality which are cumbersome and expensive, costing up to US$6000, and thus considered inappropriate for use in our system.

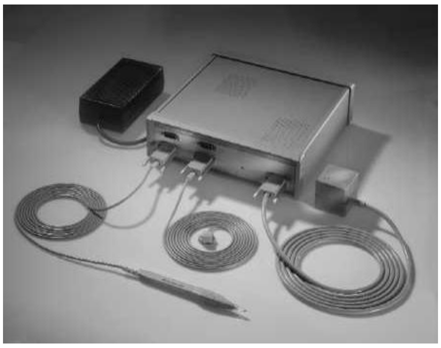


Figure 2.3: Polhemus FasTrak

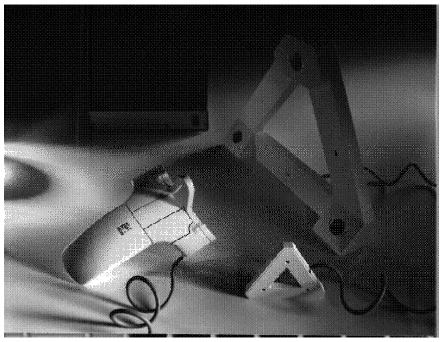


Figure 2.4: Logitech 3D Mouse

Figure 2.5: Logitech Spaceball

The *Logitech Spaceball*, *Gyration Gyromouse* and *Interlink RemotePoint* represent another category of input devices designed for personal use as a mouse replacement. The Spaceball is a device with a ball-shaped controller mounted on top. It allows users to push, pull and twist the ball in order to manipulate on-screen objects. It is designed

Figure 2.6: Gyration Gyromouse

for 3D model manipulation and provides a more natural movement for the user. The Gyromouse is based on a technology called GyroPoint that uses gyroscopes to detect angular movements of the device. These rotations can be used to control a 3D object or mouse cursor. The RemotePoint allow users to roll their thumb around a soft rubber pointing button fitted onto a handheld device. The advantages over the previous set include their ability to be wireless, more affordable and natural, although the user still interacts through an intermediary device.

Figure 2.7: Interlink RemotePoint

Figure 2.8: Origin HeadMouse

Another category of input devices tracks the position of the head of the user and moves the mouse cursor on the display correspondingly. These are primarily developed to provide full mouse control to people who cannot use their hands but have good head control. Devices in this category consist of the *Synapse Head Tracking Device, Origin Instruments HeadMouse* and *NaturalPoint Smart-Nav.*

Figure 2.9: Natural Point Smart-Nav

Eye tracking devices follow the movement of the pupils and can in principle be used for cursor control [17]. This technology is beneficial within a confined, controlled space but there is currently no evidence to support its use on a large display. Another promising input system is voice-recognition and although recognition accuracy is improving, these systems are primarily command-based [14],

Figure 2.10: Eye tracking device

thus indirect.

Touch screens provide an excellent solution to the problem of direct interaction and have high precision [18] but again, they do not scale well to large displays as manufacturing of such large touch display is not feasible (even with technology such as *DiamondTouch* [19] or *Smartskin* [20]). Even if it were possible, the user cannot reach the top of the display nor pace across from side to side. The same can be said with *MimioMouse* [21] which uses a pen like device to allow the user to press against the wall at a specific location on the display.

Figure 2.11: Smartskin on a Touch screen

Figure 2.12: Mimio Mouse

Recent inventions such as the handheld *Personal Digital Assistant* (PDA) and *graphics tablets* use a stylus device, such as a specialised pen, to perform input. Such devices are only useful for personal use, where they allow direct interaction. However they are still indirect when used with a large display because the user performs input on the handheld device and receives feedback on the large display on the wall.

Figure 2.13: Wacom tablet

Figure 2.14: Compaq PDA

One device that is attracting an increasing amount of research is the laser pointer. These have the advantages of mobility, direct interaction, and being comparably inexpensive with the notable disadvantages of lag and instability with the human hand.

Many studies into these systems have been carried out. Dwelling is a popular technique for interaction [22], and Olsen investigated the effect of lag with this method [23], which led to a discussion on the use of visible and invisible laser pointers [24]. The problem of hand jitter was presented in [25, 26].

Figure 2.15: Laser Pointer

As can be seen, some of the devices discussed above required a cable and some do not provide interaction at a distance such as the touch screens. Of the ones that provide direct interaction, some have limited usable area such as the industrial strengths tracking devices. Others such as the NaturalPoint SmartNav or a laser pointer based system require video based tracking. Such tracking devices require a line-of-sight between the tracking camera and the device or an indicator produced onscreen. Occlusion can be a problem as shown in figure 2.16. This restricts the position of the user and increases annoyance, it is therefore unsatisfactory.

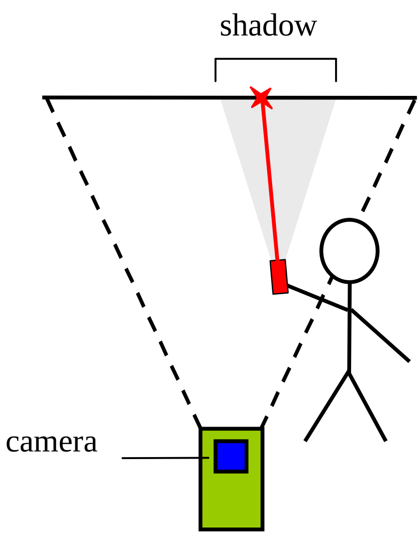


Figure 2.16: A user occluding the tracking camera and the indicator to track.

Devices such as the Gyromouse and the RemotePoint do not require such tracking. A radio link is used to transmit data from the device to a receiver. However, the major disadvantage if these devices is that they are indirect. Users cannot use these devices to directly point at a location, they need to learn to use the device and use it so that the onscreen cursor can move to the desired location.

It would be good if a system could be developed to make use of the advantage of having direct interaction with a laser pointer and the advantage of using radio link to transmit the data. Such device would indeed be advantageous since it provides a perfect solution to direct interaction with large display at a distance.

Chapter 3

Research Methods and Techniques

The proposed solution is illustrated in figure 3.1. It is setup to be used in a presentation environment. Initially, a computer provides some data to be projected onto the large display using a data projector. A user will then be using our pointing device to interact with the display. The pointing device will have a transmitter at either end which transmits radio signals and picked up by a stationary receiver. This enables the system to determine the exact location of the transmitters T1 and T2 (figure 3.2) in 3D space (x,y,z coordinate). Using these two coordinates, we can extend a straight line from these two points to the display which we refer to as *I* - the inferred coordinates of the pointing device. It indicates the position where the user wants to point to. The system then computes the position on the display in terms of the coordinate of the display. A visible red laser pointer could be attached to the device so that visual feedback can be supplied to the user so that they know exactly where they are aiming at.

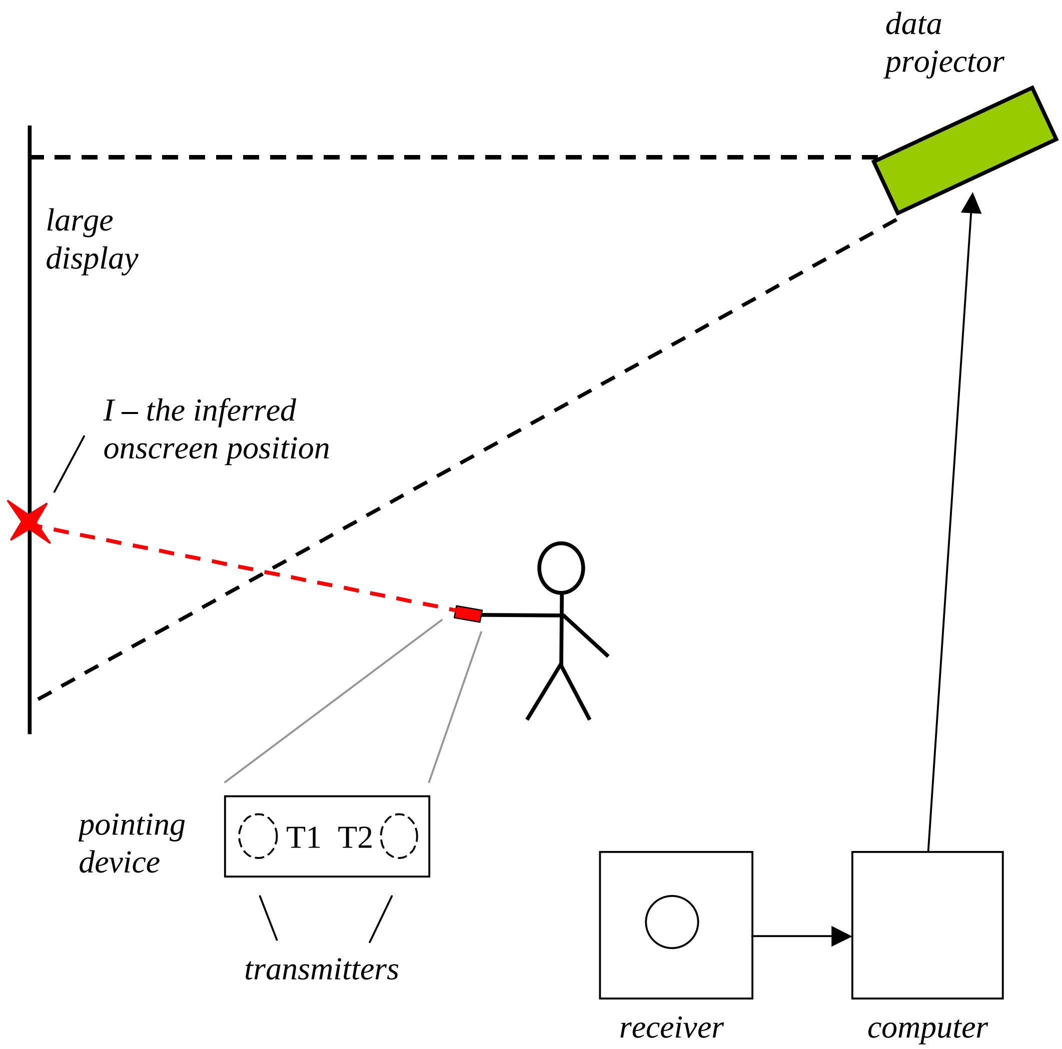


Figure 3.1: A diagram of the proposed system

Figure 3.2: I – the inferred pointer location, T1 – the front transmitter, T2 – the back transmitter

Most pointing devices require a button for clicking so that items onscreen can be selected. However, with our proposed pointing device, no such button is available. Kirstein and Muller [22] suggested the use of the on/off switch to act as a button, where the pointer is switched off momentarily to signify a click. However, as suggested in [26] when the button is released, the beam often moves away from the target before it goes off, therefore it does not provide a good indication of the intended selection. Dwell clicking is a possible candidate for replacing a button click [23]. However, the major disadvantage of this technique is that it takes at least 2 seconds to make a dwell selection [26]. As suggested by [2] a software solution can be used as an alternative which provides greater flexibility and can be applied to a whole range of devices. Winograd and Guimbretiere [27] uses a concept called *action bar* which can be activated by sweeping the pointer across them (figure 3.3). While

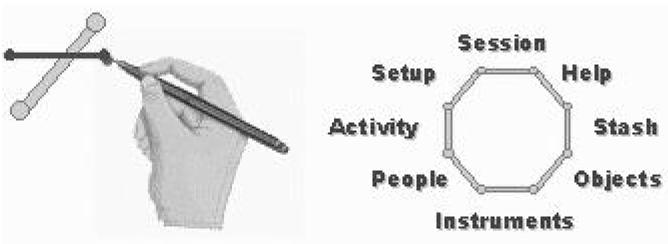
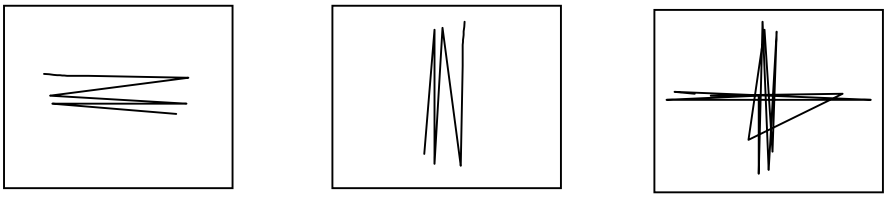


Figure 3.3: action bar

Sukthankar *et al* [28] proposed that if the pointing device falls within an *active region* (a bounding box of an object), the action associated with that region will be triggered. A circling gesture can also be used to select an object [2], however it is sometimes difficult to circle an object because of the size of the object. In this project, we have proposed an easier solution to circling. The idea is to select an object when a "zigzag" gesture is made by the pointer. This is easier for most users because it does not require any special gesture but relying on their natural hand movement. A zigzag movement can be done with an elaborate shaking of the hand as shown in figure 3.4. Also to make it more robust, gestures are allowed to go outside of the bounding box of the object, but must finish off the gesture inside the bounding box. The zigzag can go in any direction but should be consistent in one single gesture. However, one natural problem is that as the hand of the user is unsteady [26], objects will be



(a) (b) (c)

Figure 3.4: (a),(b) correct gesture, (c) incorrect gesture

selected undesirably. This problem is solved if we can make sure that the zigzag gesture has a minimum size, we can then be sure that it is not accidental.

The main concern with this project is whether a 3D coordinate can be generated using a single radio transmitter and a single radio receiver. We will need to consult with experts from the field of mathematics and physics and determine whether such system is possible. If it is proven unachievable, we will need to rely on another technology. Ultrasound is a medium used by [29] used to detect a unit called “Bats” in 3D space, so perhaps this would be a possible alternative to radio signal.

Another issue is whether the transmitter will be small enough to fit into a handheld shape device. If this is not the case, we will need to determine the minimum size of the device so that it is sufficient for the transmitter to be fitted inside. There is then the question of whether the user is comfortable holding such a device around when presenting.

The system will be tested out in a presentation environment as mentioned previously. Users will be required to test the effectiveness of this system. For efficiency, we will set up a target on the presentation, for example a rectangular box in a particular location on the slide, and see how long it takes the users to acquire the target using our device. This can also test for any latency issues.

For the evaluation of the gesture, users will be asked to make the zigzag gesture in a way that they would feel comfortable with. A variety of shapes and sizes will be tested to see if there is a correlation between them.

Chapter 4

Research Plan and Schedule

March 2003 – July 2003 (6 months)	Literature review covering input devices, interaction techniques and gestures.
August 2003 – November 2003 (4 months)	Seek advice for the radio link problem and see if it is feasible. Assess alternative medium or methods for getting the coordinates in 3D space.
December 2004 – May 2004 (6 months)	Implement the system using radio wave and devising the device.
June 2004 – August 2004 (3 months)	User evaluation of the system in a presentation environment.
September 2004 – February 2005 (6 months)	Implement the zigzag gesture and determine other possible gestures.
March 2005 – May 2005 (3 months)	User evaluation of the system in the presentation environment using both the pointing device and the gestures.
June 2005 – February 2006 (9 months)	Writing up the thesis and complete any unfinished work.

References

[1] Engelbart, D. and English, W., "A Research Center for Augmenting Human Intellect", *Reprinted in ACM SIGGRAPH Video Review*, *1994.*, 1968. 106

[2] Cheng, K., "Direct Interaction with Large-Scale Display Systems using Laser Pointers", Honours Thesis, The University of Sydney, 2002.

[3] Robertson, G. G. and Mackinlay, J. D., "The Document Lens", In *Proceedings of UIST '93*, pages 101-108.

[4] Furnas, G.W., "Generalized fisheye views," *SIGCHI'86 Conference on Human Factors in Computing Systems*, ACM Press, p. 16-23.

[5] Lamping, J., Rao, R., and Pirolli, P. "A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies", *Proceedings of CHI '95*, ACM Press, p.401-408.

[6] Chuah, M. C., Roth, S. F., Mattis, J. and Kolojejchick, J., "SDM: Selective Dynamic Manipulation of Visualizations", *Proceedings of UIST'95 Conference,* p.15-17.

[7] Chen, X. and Davis, J., “LumiPoint: Multi-User Laser-Based Interaction on Large Tiled Displays”, Technical report, Stanford University, 2001.

[8] Bier, E. A., Stone, M. C., Pier, K, Buxton, W., & DeRose, T. D., “Toolglass and magic lenses: The see-through interface”, *Proceedings of SIGGRAPH '93,* p.73- 80.

[9] Guimbretière, F. and Winograd, T., “FlowMenu: Combining Command, Text, and Data Entry”, *Proceedings of the 13th annual ACM symposium on User interface software and technology 2000*, p.213-216.

[10] Cotter, C., “Turtle Graphics Interface for REDUCE Version 3”, 1998

[11] Shneiderman, B., “Direct Manipulation: A Step Beyond Programming Languages”, *IEEE Computer*, August 1983, p.57-69.

[12] *MacDraw Manual*. Apple Computer, Inc., 1984.

[13] Butterworth, J., Davidson, A., Hench, S. and Olano, T. M, “3DM: A Three Dimensional Modeler Using a Head-Mounted Display", Proceedings of ACM Symposium on Interactive 3D Graphics, p.135-139, 1992.

[14] Jacob, R. J., “Human-Computer Interaction: Input Devices”, *ACM Computing Surveys*, 1996, 28(1): p.177-179.

[15] Myers, B. A., "A Brief History of Human Computer Interaction Technology", ACM Interactions, 1998, 5(2), p. 44-54.

[16] Keefe, D., Acevado, D., Moscovish, T., Laidlaw, D., and LaViola, J., “CavePainting: A Fully Immersive 3D Artistic Medium and Interactive Experience”, *Proceedings of the 2001 Symposium on Interactive 3D Graphics*, pages 85-93.

[17] Sibert, L. E. and Jacob, R. J. K., “Evaluation of Eye Gaze Interaction”, *Proceedings of the CHI 2000*, pages 281-288.

[18] Sears, A., Shneiderman, B., "High Precision Touchscreens: Design Strategies and Comparisons with a Mouse", *International Journal of Man-Machine Studies, 34 (4),* 1991, p. 593-613.

[19] Dietz, P.H. and Leigh, D.L., "DiamondTouch: A Multi-User Touch Technology", *ACM Symposium on User Interface Software and Technology (UIST)*, 2001, p.219-226.

[20] Rekimoto, J., "SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces", *ACM CHI 2002 Conference on Human Factors in Computing Systems*, p.113-120.

[21] MimioMouse, http://www.mimio.com/

[22] Kirstein, C. and Muller, H., "Interaction with a projection screen using a camera-tracked laser pointer," in *Multimedia Modeling; MMM '98 Proceedings*, 1998, p.191-192.

[23] Olsen Jr, D.R. and Nielsen, T., "Laser Pointer Interaction", *ACM CHI'2001 Conference Proceedings: Human Factors in Computing Systems*, 2001, Seattle, WA: p.17-22.

[24] Cavens, D., Vogt, F., Fels, S., and Meitner, M., "Interacting with the big screen: pointers to ponder", *ACM SIGCHI '02 Conference on Human Factors in Computing System, Companion Proceedings*, p.181-182.

[25] Myers, B.A., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R., and Long, A. C., "Interacting At a Distance: Measuring the Performance of Laser Pointers and Other Devices", *ACM CHI 2002 Conference on Human Factors in Computing Systems*, p.33-40.

[26] Peck, C., "Useful Parameters for the Design of Laser Pointer Interaction Techniques", in *Proceedings of CHI 2001*, p.461-462.

[27] Winograd, T. and Guimbretirere, F., “Visual Instruments for an interactive Mural”, *ACM SIGCHI CHI99 Extended Abstracts*, p234-235.

[28] Sukthankar, R., Stockton, R. and Mullin, M., “Self-Calibrating Camera-Assisted Presentation Interface”, *Proceedings of International Conference on Automation, Control, Robotics and Computer Vision, 2000*

[29] Harter, A., Hopper A., Steggles, P., Ward, A. and Webster, P., "The anatomy of a context-aware application," in Proceedings of MOBICOM 1999, p.59-68.