

# ThinSight: A Thin Form-Factor Interactive Surface Technology

By Shahram Izadi, Steve Hodges, Alex Butler, Darren West, Alban Rrustemi, Mike Molloy and William Buxton

## ABSTRACT

**ThinSight is a thin form-factor interactive surface technology based on optical sensors embedded inside a regular liquid crystal display (LCD). These augment the display with the ability to sense a variety of objects near the surface, including fingertips and hands, to enable multitouch interaction. Optical sensing also allows other physical items to be detected, allowing interactions using various tangible objects. A major advantage of ThinSight over existing camera and projector-based systems is its compact form-factor, making it easier to deploy in a variety of settings. We describe how the ThinSight hardware is embedded behind a regular LCD, allowing sensing without degradation of display capability, and illustrate the capabilities of our system through a number of proof-of-concept hardware prototypes and applications.**

## 1. INTRODUCTION

Touch input using a single point of contact with a display is a natural and established technique for human computer interaction. Research over the past decades,<sup>3</sup> and more recently products such as the iPhone and Microsoft Surface, have shown the novel and exciting interaction techniques and applications possible if multiple simultaneous touch points can be detected.

Various technologies have been proposed for multitouch sensing in this way, some of which extend to detection of physical objects in addition to fingertips. Systems based on optical sensing have proven to be particularly powerful in the richness of data captured and the flexibility they can provide. As yet, however, such optical systems have predominately been based on cameras and projectors and require a large optical path in front of or behind the display. This typically results in relatively bulky systems—something that can impact adoption in many real-world scenarios. While capacitive overlay technologies, such as those in the iPhone and the Dell XT Tablet PC, can support thin form-factor multitouch, they are limited to sensing only fingertips.

ThinSight is a novel interactive surface technology which is based on optical sensors integrated into a thin form-factor LCD. It is capable of imaging multiple fingertips, whole hands, and other objects near the display surface as shown in Figure 1. The system is based upon custom hardware embedded behind an LCD, and uses infrared (IR) light for sensing without degradation of display capability.

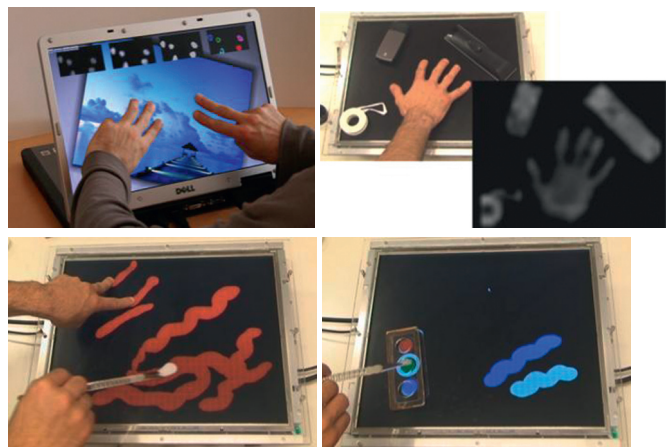
In this article we describe the ThinSight electronics and the modified LCD construction which results. We present two prototype systems we have developed: a multitouch laptop and a touch-and-tangible tabletop (both shown in Figure 1). These

systems generate rich sensor data which can be processed using established computer vision techniques to prototype a wide range of interactive surface applications.

As shown in Figure 1, the shapes of many physical objects, including fingers, brushes, dials, and so forth, can be “seen” when they are near the display, allowing them to enhance multitouch interactions. Furthermore, ThinSight allows interactions close-up or at a distance using active IR pointing devices, such as styluses, and enables IR-based communication through the display with other electronic devices.

We believe that ThinSight provides a glimpse of a future where display technologies such as LCDs and organic light emitting diodes (OLEDs) will cheaply incorporate optical sensing pixels alongside red, green and blue (RGB) pixels in

**Figure 1. ThinSight brings the novel capabilities of surface computing to thin displays. Top left: photo manipulation using multiple fingers on a laptop prototype (note the screen has been reversed in the style of a Tablet PC). Top right: a hand, mobile phone, remote control and reel of tape placed on a tabletop ThinSight prototype, with corresponding sensor data far right. Note how all the objects are imaged through the display, potentially allowing not only multitouch but tangible input. Bottom left and right: an example of how such sensing can be used to support digital painting using multiple fingertips, a real brush and a tangible palette to change paint colors.**



Original versions of this paper appeared in *Proceedings of the 2007 ACM Symposium on User Interface Software and Technology* as “ThinSight: Versatile Multi-touch Sensing for Thin Form-factor Displays” and in *Proceedings of the 2008 IEEE Workshop on Horizontal Interactive Human Computer Systems* as “Experiences with Building a Thin Form-Factor Touch and Tangible Tabletop.”

a similar manner, resulting in the widespread adoption of such surface technologies.

## 2. OVERVIEW OF OPERATION

### 2.1. Imaging through an LCD using IR light

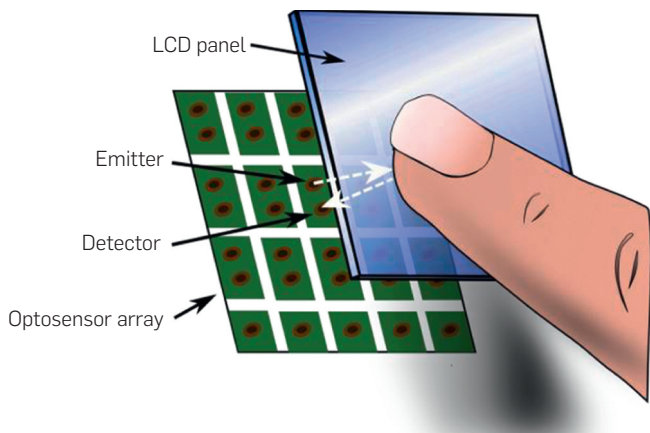
A key element in the construction of ThinSight is a device known as a retro-reflective optosensor. This is a sensing element which contains two components: a light emitter and an optically isolated light detector. It is therefore capable of both emitting light and, at the same time, detecting the intensity of incident light. If a reflective object is placed in front of the optosensor, some of the emitted light will be reflected back and will therefore be detected.

ThinSight is based around a 2D grid of retro-reflective optosensors which are placed behind an LCD panel. Each optosensor emits light that passes right through the entire panel. Any reflective object in front of the display (such as a fingertip) will reflect a fraction of the light back, and this can be detected. Figure 2 depicts this arrangement. By using a suitably spaced grid of retro-reflective optosensors distributed uniformly behind the display it is therefore possible to detect any number of fingertips on the display surface. The raw data generated is essentially a low resolution grayscale “image” of what can be seen through the display, which can be processed using computer vision techniques to support touch and other input.

A critical aspect of ThinSight is the use of retro-reflective sensors that operate in the infrared part of the spectrum, for three main reasons:

- Although IR light is attenuated by the layers in the LCD panel, some still passes through the display.<sup>5</sup> This is largely unaffected by the displayed image.
- A human fingertip typically reflects around 20% of incident IR light and is therefore a quite passable “reflective object.”
- IR light is not visible to the user, and therefore does not detract from the image being displayed on the panel.

**Figure 2. The basic principle of ThinSight. An array of retro-reflective optosensors is placed behind an LCD. Each of these contains two elements: an emitter which shines IR light through the panel; and a detector which picks up any light reflected by objects such as fingertips in front of the screen.**



### 2.2. Further features of ThinSight

ThinSight is not limited to detecting fingertips in contact with the display; any suitably reflective object will cause IR light to reflect back and will therefore generate a “silhouette.” Not only can this be used to determine the location of the object on the display, but also its orientation and shape, within the limits of sensing resolution. Furthermore, the underside of an object may be augmented with a visual mark—a barcode of sorts—to aid identification.

In addition to the detection of passive objects via their shape or some kind of barcode, it is also possible to embed a very small infrared transmitter into an object. In this way, the object can transmit a code representing its identity, its state, or some other information, and this data transmission can be picked up by the IR detectors built into ThinSight. Indeed, ThinSight naturally supports bidirectional IR-based data transfer with nearby electronic devices such as smartphones and PDAs. Data can be transmitted from the display to a device by modulating the IR light emitted. With a large display, it is possible to support several simultaneous bidirectional communication channels in a spatially multiplexed fashion.

Finally, a device which emits a collimated beam of IR light may be used as a pointing device, either close to the display surface like a stylus, or from some distance. Such a pointing device could be used to support gestures for new forms of interaction with a single display or with multiple displays. Multiple pointing devices could be differentiated by modulating the light generated by each.

## 3. THE THINSIGHT HARDWARE

### 3.1. The sensing electronics

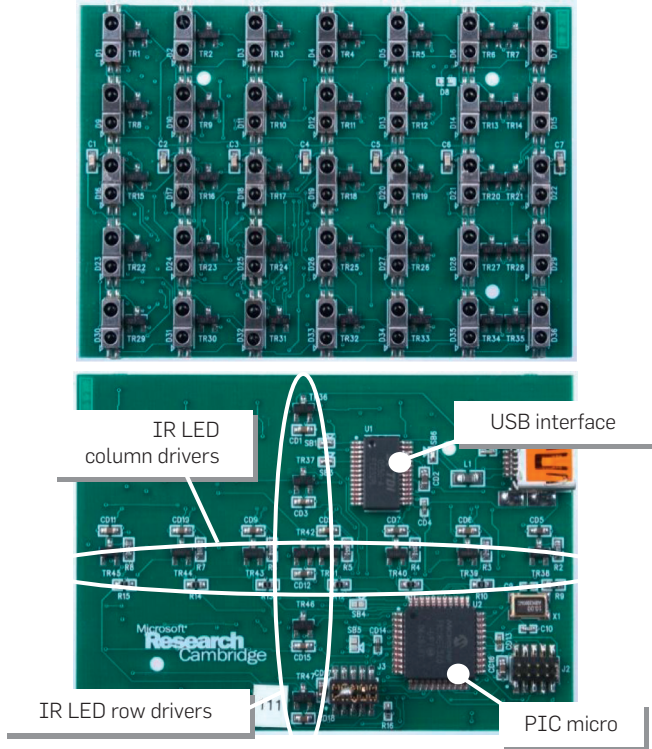
The prototype ThinSight circuit board depicted in Figure 3 uses Avago HSDL-9100 retro-reflective infrared sensors. These devices are especially designed for proximity sensing—an IR LED emits infrared light and an IR photodiode generates a photocurrent which varies with the amount of incident light. Both emitter and detector have a center wavelength of 940 nm.

A  $7 \times 5$  grid of these HSDL-9100 devices on a regular 10 mm pitch is mounted on custom-made  $70 \times 50$  mm 4-layer printed circuit board (PCB). Multiple PCBs can be tiled together to support larger sensing areas. The IR detectors are interfaced directly with digital input/output lines on a PIC18LF4520 microcontroller.

The PIC firmware collects data from one row of detectors at a time to construct a “frame” of data which is then transmitted to the PC over USB via a virtual COM port. To connect multiple PCBs to the same PC, they must be synchronized to ensure that IR emitted by a row of devices on one PCB does not adversely affect scanning on a neighboring PCB. In our prototype we achieve this using frame and row synchronization signals which are generated by one of the PCBs (the designated “master”) and detected by the others (“slaves”).

Note that more information on the hardware can be found in the full research publications.<sup>7,10</sup>

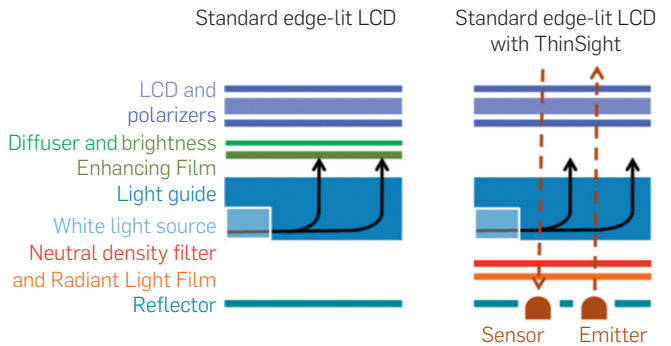
**Figure 3. Top:** the front side of the sensor PCB showing the 7×5 array of IR optosensors. The transistors that enable each detector are visible to the right of each optosensor. **Bottom:** the back of the sensor PCB has little more than a PIC microcontroller, a USB interface and FETs to drive the rows and columns of IR emitting LEDs. Three such PCBs are used in our ThinSight laptop while there are thirty in the tabletop prototype.



### 3.2. LCD technology overview

To understand how the ThinSight hardware is integrated into a display panel, it is useful to understand the construction and operation of a typical LCD. An LCD panel is made up of a stack of optical components as shown in Figure 4. At the front of the panel is a thin layer of liquid crystal material which is sandwiched between two polarizers. The polarizers are orthogonal to each other, which means that any light which passes through the first will naturally be blocked by the second, resulting in dark pixels. However, if a voltage is applied across the liquid crystal material at a certain pixel location, the polarization of light incident on that pixel is twisted through 90° as it passes through the crystal structure. As a result it emerges from the crystal with the correct polarization to pass through the second polarizer. Typically, white light is shone through the panel from behind by a backlight and red, green, and blue filters are used to create a color display. In order to achieve a low profile construction while maintaining uniform lighting across the entire display and keeping cost down, the backlight is often a large “light guide” in the form of a clear acrylic sheet which sits behind the entire LCD and which is edge-lit from one or more sides. The light source is often a cold cathode fluorescent tube or an array of white LEDs. To maximize the efficiency and uniformity of the lighting, additional layers of material may

**Figure 4. Typical LCD edge-lit architecture shown left.** The LCD comprises a stack of optical elements. A white light source is typically located along one or two edges at the back of the panel. A white reflector and transparent light guide direct the light toward the front of the panel. The films help scatter this light uniformly and enhance brightness. However, they also cause excessive attenuation of IR light. **In ThinSight, shown right,** the films are substituted and placed behind the light guide to minimize attenuation and also reduce noise caused by LCD flexing upon touch. The sensors and emitters are placed at the bottom of the resulting stack, aligned with holes cut in the reflector.



be placed between the light guide and the LCD. Brightness enhancing film (BEF) “recycles” visible light at suboptimal angles and polarizations and a diffuser smoothes out any local nonuniformities in light intensity.

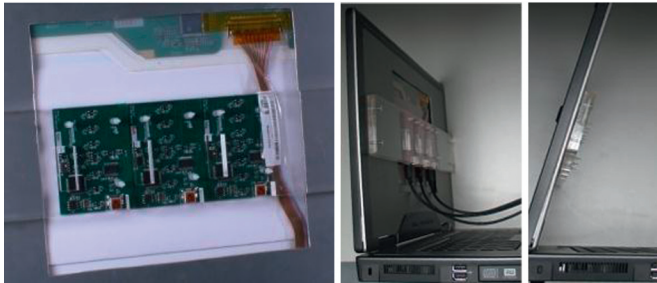
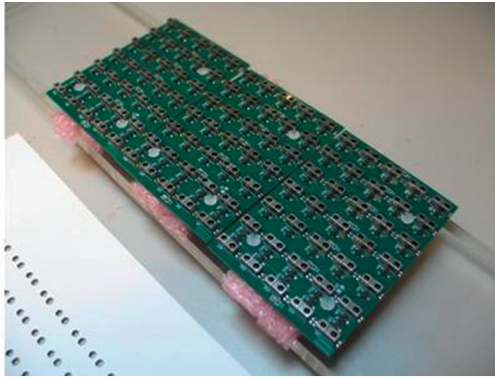
### 3.3. Integration with an LCD panel

We constructed our ThinSight prototypes using a variety of desktop and laptop LCD panels, ranging from 17” to 21”. Two of these are shown in Figures 5 and 6. Up to 30 PCBs were tiled to support sensing across the entire surface. In instances where large numbers of PCBs were tiled, a custom hub circuit based on an FPGA was designed to collect and aggregate the raw data captured from a number of tiled sensors and transfer this to the PC using a single USB channel. These tiled PCBs are mounted directly behind the light guide. To ensure that the cold cathode does not cause any stray IR light to emanate from the acrylic light guide, we placed a narrow piece of IR-blocking film between it and the backlight. We cut small holes in the white reflector behind the light guide to coincide with the location of every IR emitting and detecting element.

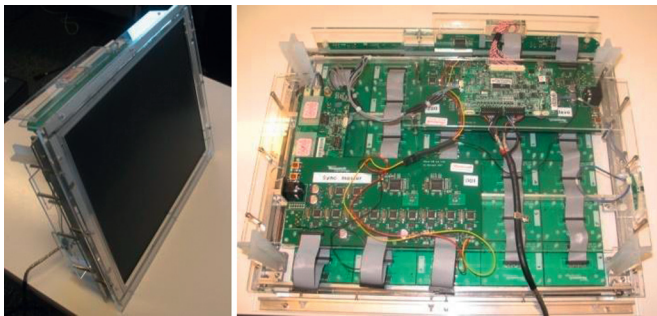
During our experiments we found that the combination of the diffuser and BEF in an LCD panel typically caused excessive attenuation of the IR signal. However, removing these materials degrades the displayed image significantly: without BEF the brightness and contrast of the displayed image is reduced unacceptably; without a diffuser the image appears to “float” in front of the backlight and at the same time the position of the IR emitters and detectors can be seen in the form of an array of faint dots across the entire display.

To completely hide the IR emitters and detectors we required a material that lets IR pass through it but not visible light, so that the optosensors could not be seen but would operate normally. The traditional solution would be

**Figure 5. Our laptop prototype. Top: Three PCBs are tiled together and mounted on an acrylic plate, to give a total of 105 sensing pixels. Holes are also cut in the white reflector shown on the far left. Bottom left: an aperture is cut in the laptop lid to allow the PCBs to be mounted behind the LCD. This provides sensing across the center of the laptop screen. Bottom right: side views of the prototype—note the display has been reversed on its hinges in the style of a Tablet PC.**



**Figure 6. The ThinSight tabletop hardware as viewed from the side and behind. Thirty PCBs (in a 5×6 grid) are tiled with columns interconnected with ribbon cable and attached to a hub board for aggregating data and inter-tile communication. This provides a total of 1050 discrete sensing pixels across the entire surface.**



to use what is referred to as a “cold mirror.” Unfortunately these are made using a glass substrate which means they are expensive, rigid and fragile and we were unable to source a cold mirror large enough to cover the entire tabletop display. We experimented with many alternative materials including tracing paper, acetate sheets coated in emulsion paint, spray-on frosting, thin sheets of white polythene and mylar. Most of these are unsuitable either because of

a lack of IR transparency or because the optosensors can be seen through them to some extent. The solution we settled on was the use of Radiant Light Film by 3M (part number CM500), which largely lets IR light pass through while reflecting visible light without the disadvantages of a true cold mirror. This was combined with the use of a grade “0” neutral density filter, a visually opaque but IR transparent diffuser, to even out the distribution rear illumination and at the same time prevent the “floating” effect. Applying the Radiant Light Film carefully is critical since minor imperfections (e.g. wrinkles or bubbles) are highly visible to the user—thus we laminated it onto a thin PET carrier. One final modification to the LCD construction was to deploy these films *behind* the light guide to further improve the optical properties. The resulting LCD layer stack-up is depicted in Figure 4 right.

Most LCD panels are not constructed to resist physical pressure, and any distortion which results from touch interactions typically causes internal IR reflection resulting in “flare.” Placing the Radiant Light Film and neutral density filter behind the light guide improves this situation, and we also reinforced the ThinSight unit using several lengths of extruded aluminum section running directly behind the LCD.

## 4. THINSIGHT IN OPERATION

### 4.1. Processing the raw sensor data

Each value read from an individual IR detector is defined as an integer representing the intensity of incident light. These sensor values are streamed to the PC via USB where the raw data undergoes several simple processing and filtering steps in order to generate an IR image that can be used to detect objects near the surface. Once this image is generated, established image processing techniques can be applied in order to determine coordinates of fingers, recognize hand gestures, and identify object shapes.

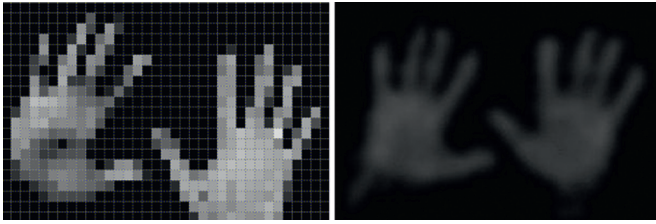
Variations between optosensors due to manufacturing and assembly tolerances result in a range of different values across the display even without the presence of objects on the display surface. To make the sensor image uniform and the presence of additional incident light (reflected from nearby objects) more apparent, we subtract a “background” frame captured when no objects are present, and normalize relative to the image generated when the display is covered with a sheet of white reflective paper.

We use standard bicubic interpolation to scale up the sensor image by a predefined factor (10 in our current implementation). For the larger tabletop implementation this results in a 350 × 300 pixel image. Optionally, a Gaussian filter can be applied for further smoothing, resulting in a grayscale “depth” image as shown in Figure 7.

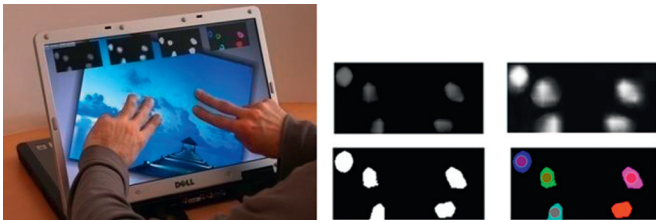
### 4.2. Seeing through the ThinSight display

The images we obtain from the prototype are quite rich, particularly given the density of the sensor array. Fingers and hands within proximity of the screen are clearly identifiable. Examples of images captured through the display are shown in Figures 1, 7 and 8.

**Figure 7.** The raw ThinSight sensor data shown left and after interpolation and smoothing right. Note that the raw image is a very low resolution, but contains enough data to generate the relatively rich image at right.



**Figure 8.** Fingertips can be sensed easily with ThinSight. Left: the user places five fingers on the display to manipulate a photo. Right: a close-up of the sensor data when fingers are positioned as shown at left. The raw sensor data is: (1) scaled-up with interpolation, (2) normalized, (3) thresholded to produce a binary image, and finally (4) processed using connected components analysis to reveal the fingertip locations.



Fingertips appear as small blobs in the image as they approach the surface, increasing in intensity as they get closer. This gives rise to the possibility of sensing both touch and hover. To date we have only implemented touch/no-touch differentiation, using thresholding. However, we can reliably and consistently detect touch to within a few millimeters for a variety of skin tones, so we believe that disambiguating hover from touch would be possible.

In addition to fingers and hands, optical sensing allows us to observe other IR reflective objects through the display. Figure 1 illustrates how the display can distinguish the shape of many reflective objects in front of the surface, including an entire hand, mobile phone, remote control, and a reel of white tape. We have found in practice that many objects reflect IR.

A logical next step is to attempt to uniquely identify objects by placement of visual codes underneath them. Such codes have been used effectively in tabletop systems such as the Microsoft Surface and various research prototypes<sup>12, 28</sup> to support tangible interaction. We have also started preliminary experiments with the use of such codes on ThinSight, see Figure 9.

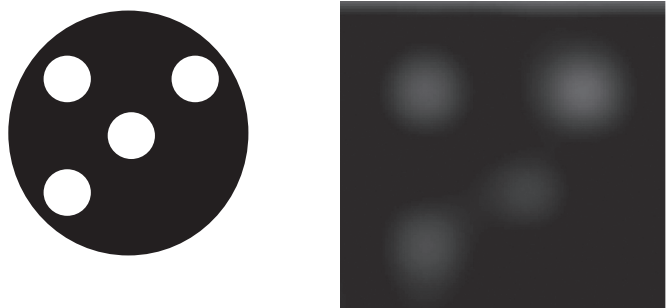
Active electronic identification schemes are also feasible. For example, cheap and small dedicated electronic units containing an IR emitter can be stuck onto or embedded inside objects that need to be identified. These emitters will produce a signal directed to a small subset of the display sensors. By emitting modulated IR it is possible to transmit a unique identifier to the display.

### 4.3. Communicating through the ThinSight display

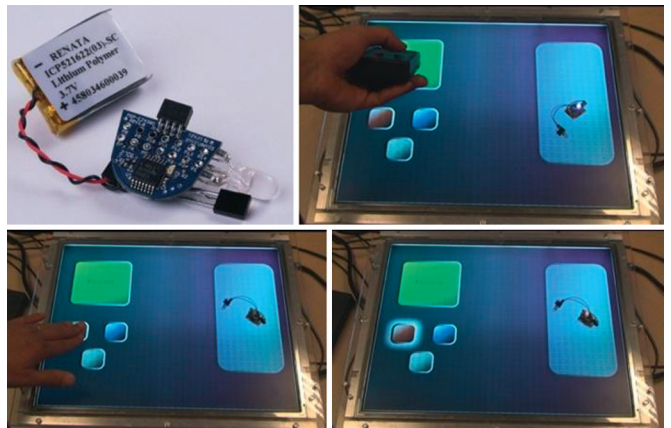
Beyond simple identification, an embedded IR transmitter also provides a basis for supporting richer bidirectional communication with the display. In theory any IR modulation scheme, such as the widely adopted IrDA standard, could be supported by ThinSight. We have implemented a DC-balanced modulation scheme which allows retro-reflective object sensing to occur *at the same time* as data transmission. This required no additions or alterations to the sensor PCB, only changes to the microcontroller firmware. To demonstrate our prototype implementation of this, we built a small embedded IR transceiver based on a low power MSP430 microcontroller, see Figure 10. We encode 3 bits of data in the IR transmitted from the ThinSight pixels to control an RGB LED fitted to the embedded receiver. When the user touches various soft buttons on the ThinSight display, this in turn transmits different 3 bit codes from ThinSight pixels to cause different colors on the embedded device to be activated.

It is theoretically possible to transmit and receive different data simultaneously using different columns on the

**Figure 9.** An example 2" diameter visual marker and the resulting ThinSight image after processing.



**Figure 10.** Using ThinSight to communicate with devices using IR. Top left: an embedded microcontroller/IR transceiver/RGB LED device. Bottom left: touching a soft button on the ThinSight display signals the RGB LED on the embedded device to turn red (bottom right). Top right: A remote control is used to signal from a distance the display which in turn sends an IR command to the RGB device to turn the LED blue.



display surface, thereby supporting spatially multiplexed bidirectional communications with multiple local devices and reception of data from remote gesturing devices. Of course, it is also possible to time multiplex communications between different devices if a suitable addressing scheme is used. We have not yet prototyped either of these multiple-device communications schemes.

#### 4.4. Interacting with ThinSight

As shown earlier in this section, it is straightforward to sense and locate multiple fingertips using ThinSight. In order to do this we threshold the processed data to produce a binary image. The connected components within this are isolated, and the center of mass of each component is calculated to generate representative  $X, Y$  coordinates of each finger. A very simple homography can then be applied to map these fingertip positions (which are relative to the sensor image) to onscreen coordinates. Major and minor axis analysis or more detailed shape analysis can be performed to determine orientation information. Robust fingertip tracking algorithms or optical flow techniques<sup>28</sup> can be employed to add stronger heuristics for recognizing gestures.

Using these established techniques, fingertips are sensed to within a few millimeters, currently at 23 frames/s. Both hover and touch can be detected, and could be disambiguated by defining appropriate thresholds. A user therefore need not apply any force to interact with the display. However, it is also possible to estimate fingertip pressure by calculating the increase in the area and intensity of the fingertip “blob” once touch has been detected.

Figure 1 shows two simple applications developed using ThinSight. A simple photo application allows multiple images to be translated, rotated, and scaled using established multifinger manipulation gestures. We use distance and angle between touch points to compute scale factor and rotation deltas. To demonstrate some of the capabilities of ThinSight beyond just multitouch, we have built an example paint application that allows users to paint directly on the surface using both fingertips and real paint brushes. The latter works because ThinSight can detect the brushes’ white bristles which reflect IR. The paint application also supports a more sophisticated scenario where an artist’s palette is placed on the display surface. Although this is visibly transparent, it has an IR reflective marker on the underside which allows it to be detected by ThinSight, whereupon a range of paint colors are rendered underneath it. The user can change color by “dipping” either a fingertip or a brush into the appropriate well in the palette. We identify the presence of this object using a simple ellipse matching algorithm which distinguishes the larger palette from smaller touch point “blobs” in the sensor image. Despite the limited resolution of ThinSight, it is possible to differentiate a number of different objects using simple silhouette shape information.

## 5. DISCUSSION AND FUTURE WORK

We believe that the prototype presented in this article is an interesting proof-of-concept of a new approach to multi-touch and tangible sensing for thin displays. We have already

described some of its potential; here we discuss a number of additional observations and ideas which came to light during the work.

### 5.1. Fidelity of sensing

The original aim of this project was simply to detect fingertips to enable multi-touch-based direct manipulation. However, despite the low resolution of the raw sensor data, we still detect quite sophisticated object images. Very small objects do currently “disappear” on occasion when they are midway between optosensors. However, we have a number of ideas for improving the fidelity further, both to support smaller objects and to make object and visual marker identification more practical. An obvious solution is to increase the density of the optosensors, or at least the density of IR detectors. Another idea is to measure the amount of reflected light under different lighting conditions—for example, simultaneously emitting light from neighboring sensors is likely to cause enough reflection to detect smaller objects.

### 5.2. Frame rate

In informal trials of ThinSight for a direct manipulation task, we found that the current frame rate was reasonably acceptable to users. However, a higher frame rate would not only produce a more responsive UI which will be important for some applications, but would make temporal filtering more practical thereby reducing noise and improving sub-pixel accuracy. It would also be possible to sample each detector under a number of different illumination conditions as described above, which we believe would increase fidelity of operation.

### 5.3. Robustness to lighting conditions

The retro-reflective nature of operation of ThinSight combined with the use of background substitution seems to give reliable operation in a variety of lighting conditions, including an office environment with some ambient sunlight. One common approach to mitigating any negative effects of ambient light, which we could explore if necessary, is to emit modulated IR and to ignore any nonmodulated offset in the detected signal.

### 5.4. Power consumption

The biggest contributor to power consumption in ThinSight is emission of IR light; because the signal is attenuated in both directions as it passes through the layers of the LCD panel, a high intensity emission is required. For mobile devices, where power consumption is an issue, we have ideas for improvements. We believe it is possible to enhance the IR transmission properties of an LCD panel by optimizing the materials used in its construction for this purpose—something which is not currently done. In addition, it may be possible to keep track of object and fingertip positions, and limit the most frequent IR emissions to those areas. The rest of the display would be scanned less frequently (e.g. at 2–3 frames/s) to detect new touch points.

One of the main ways we feel we can improve on power consumption and fidelity of sensing is to use a more

sophisticated IR illumination scheme. We have been experimenting with using an acrylic overlay *on top* of the LCD and using IR LEDs for edge illumination. This would allow us to sense multiple touch points using standard Frustrated Total Internal Reflection (FTIR),<sup>5</sup> but not objects. We have, however, also experimented with a material called Endlighten which allows this FTIR scheme to be extended to diffuse illumination, allowing both multitouch and object sensing with far fewer IR emitters than our current setup. The overlay can also serve the dual purpose of protecting the LCD from flexing under touch.

## 6. RELATED WORK

The area of interactive surfaces has gained particular attention recently following the advent of the iPhone and Microsoft Surface. However, it is a field with over two decades of history.<sup>3</sup> Despite this sustained interest there has been an evident lack of off-the-shelf solutions for detecting multiple fingers and/or objects on a display surface. Here, we summarize the relevant research in these areas and describe the few commercially available systems.

### 6.1. Camera-based systems

One approach to detecting multitouch and tangible input is to use a video camera placed in front of or above the surface, and apply computer vision algorithms for sensing. Early seminal work includes Krueger's VideoDesk<sup>13</sup> and the DigitalDesk,<sup>26</sup> which use dwell time and a microphone (respectively) to detect when a user is actually touching the surface. More recently, the Visual Touchpad<sup>17</sup> and C-Slate<sup>9</sup> use a stereo camera placed above the display to more accurately detect touch. The disparity between the image pairs determines the height of fingers above the surface. PlayAnywhere<sup>28</sup> introduces a number of additional image processing techniques for front-projected vision-based systems, including a shadow-based touch detection algorithm, a novel visual bar code scheme, paper tracking, and an optical flow algorithm for bimanual interaction.

Camera-based systems such as those described above obviously require direct line-of-sight to the objects being sensed which in some cases can restrict usage scenarios. Occlusion problems are mitigated in PlayAnywhere by mounting the camera off-axis. A natural progression is to mount the camera *behind* the display. HoloWall<sup>18</sup> uses IR illuminant and a camera equipped with an IR pass filter behind a diffusive projection panel to detect hands and other IR-reflective objects in front of it. The system can accurately determine the contact areas by simply thresholding the infrared image. TouchLight<sup>27</sup> uses rear-projection onto a holographic screen, which is also illuminated from behind with IR light. A number of multitouch application scenarios are enabled including high-resolution imaging capabilities. Han<sup>5</sup> describes a straightforward yet powerful technique for enabling high-resolution multitouch sensing on rear-projected surfaces based on FTIR. Compelling multitouch applications have been demonstrated using this technique. The Smart Table<sup>22</sup> uses this same FTIR technique in a tabletop form factor.

The Microsoft Surface and ReacTable<sup>12</sup> also use rear-projection, IR illuminant and a rear mounted IR camera to monitor fingertips, this time in a horizontal tabletop form-factor. These systems also detect and identify objects with IR-reflective markers on their surface.

The rich data generated by camera-based systems provides extreme flexibility. However, as Wilson discusses<sup>28</sup> this flexibility comes at a cost, including the computational demands of processing high resolution images, susceptibility to adverse lighting conditions and problems of motion blur. However, perhaps more importantly, these systems require the camera to be placed at some distance from the display to capture the entire scene, limiting their portability, practicality and introducing a setup and calibration cost.

### 6.2. Opaque embedded sensing

Despite the power of camera-based systems, the associated drawbacks outlined above have resulted in a number of parallel research efforts to develop a non-vision-based multitouch display. One approach is to embed a multitouch sensor of some kind behind a surface that can have an image projected onto it. A natural technology for this is capacitive sensing, where the capacitive coupling to ground introduced by a fingertip is detected, typically by monitoring the rate of leakage of charge away from conductive plates or wires mounted behind the display surface.

Some manufacturers such as Logitech and Apple have enhanced the standard laptop-style touch pad to detect certain gestures based on more than one point of touch. However, in these systems, using more than two or three fingers typically results in ambiguities in the sensed data. This constrains the gestures they support. Lee et al.<sup>14</sup> used capacitive sensing with a number of discrete metal electrodes arranged in a matrix configuration to support multitouch over a larger area. Westerman<sup>25</sup> describes a sophisticated capacitive multitouch system which generates x-ray-like images of a hand interacting with an opaque sensing surface, which could be projected onto. A derivative of this work was commercialized by Fingerworks.

DiamondTouch<sup>4</sup> is composed of a grid of row and column antennas which emit signals that capacitively couple with users when they touch the surface. Users are also capacitively coupled to receivers through pads on their chairs. In this way the system can identify which antennas behind the display surface are being touched and by which user, although a user touching the surface at two points can produce ambiguities. The SmartSkin<sup>21</sup> system consists of a grid of capacitively coupled transmitting and receiving antennas. As a finger approaches an intersection point, this causes a drop in coupling which is measured to determine finger proximity. The system is capable of supporting multiple points of contact by the same user and generating images of contact regions of the hand. SmartSkin and DiamondTouch also support physical objects, but can only identify an object when a user touches it. Tactex provide another interesting example of an opaque multitouch sensor, which uses transducers to measure surface pressure at multiple touch points.<sup>23</sup>

### 6.3. Transparent overlays

The systems above share one major disadvantage: they all rely on front-projection for display. The displayed image will therefore be broken up by the user's fingers, hands and arms, which can degrade the user experience. Also, a large throw distance is typically required for projection which limits portability. Furthermore, physical objects can only be detected in limited ways, if object detection is supported at all.

One alternative approach to address some of the issues of display and portability is to use a transparent sensing overlay in conjunction with a self-contained (i.e., not projected) display such as an LCD panel. DualTouch<sup>19</sup> uses an off-the-shelf transparent resistive touch overlay to detect the position of two fingers. Such overlays typically report the average position when two fingers are touching. Assuming that one finger makes contact first and does not subsequently move, the position of a second touch point can be calculated. An extension to this is provided by Loviscach.<sup>16</sup>

The Philips Entertaible<sup>15</sup> takes a different “overlay” approach to detect up to 30 touch points. IR emitters and detectors are placed on a bezel around the screen. Breaks in the IR beams detect fingers and objects. The SMART DViT<sup>22</sup> and HP TouchSmart<sup>6</sup> utilize cameras in the corners of a bezel overlay to support sensing of two fingers or styluses. With such line of sight systems, occlusion can be an issue for sensing.

The Lemur music controller from JazzMutant<sup>11</sup> uses a proprietary resistive overlay technology to track up to 20 touch points simultaneously. More recently, Balda AG and N-Trig<sup>20</sup> have both released capacitive multitouch overlays, which have been used in the iPhone and the Dell XT, respectively. These approaches provide a robust way for sensing multiple fingers touching the surface, but do not scale to whole hand sensing or tangible objects.

### 6.4. The need for intrinsically integrated sensing

The previous sections have presented a number of multitouch display technologies. Camera-based systems produce very rich data but have a number of drawbacks. Opaque sensing systems can more accurately detect fingers and objects, but by their nature rely on projection. Transparent overlays alleviate this projection requirement, but the fidelity of sensing is reduced. It is difficult, for example, to support sensing of fingertips, hands and objects.

A potential solution which addresses all of these requirements is a class of technologies that we refer to as “intrinsically integrated” sensing. The common approach behind these is to distribute sensing across the display surface, integrating the sensors with the display elements. Hudson<sup>8</sup> reports on a prototype 0.7” monochrome display where LED pixels double up as light sensors. By operating one pixel as a sensor while its neighbors are illuminated, it is possible to detect light reflected from a fingertip close to the display. The main drawbacks are the use of visible illuminant during sensing and practicalities of using LED-based displays. SensoLED uses a similar approach with

visible light, but this time based on polymer LEDs and photodiodes. A 1” diagonal sensing polymer display has been demonstrated.<sup>2</sup>

Planar<sup>1</sup> and Toshiba<sup>24</sup> were among the first to develop LCD prototypes with integrated visible light photosensors, which can detect the shadows resulting from fingertips or styluses on the display. The photosensors and associated signal processing circuitry are integrated directly onto the LCD substrate. To illuminate fingers and other objects, either an external light source is required—impacting on the profile of the system—or the screen must uniformly emit bright visible light—which in turn will disrupt the displayed image.


The motivation for ThinSight was to build on the concept of intrinsically integrated sensing. We have extended the work above using invisible (IR) illuminant to allow simultaneous display and sensing, building on current LCD and IR technologies to make prototyping practical in the near term. Another important aspect is support for much larger thin touch-sensitive displays than is provided by intrinsically integrated solutions to date, thereby making it more practical to prototype multitouch applications.

## 7. CONCLUSION

In this article we have described a new technique for optically sensing multiple objects, including fingertips, through thin form-factor displays. Optical sensing allows rich “camera-like” data to be captured by the display and this is processed using computer vision techniques. This supports new types of human computer interfaces that exploit zero-force multi-touch and tangible interaction on thin form-factor displays such as those described in Buxton.<sup>3</sup> We have shown how this technique can be integrated with off-the-shelf LCD technology, making such interaction techniques more practical and deployable in real-world settings.

We have many ideas for potential refinements to the ThinSight hardware, firmware, and PC software. In addition to such incremental improvements, we also believe that it will be possible to transition to an integrated “sensing and display” solution which will be much more straightforward and cheaper to manufacture. An obvious approach is to incorporate optical sensors directly onto the LCD backplane, and as reported earlier early prototypes in this area are beginning to emerge.<sup>24</sup> Alternatively, polymer photodiodes may be combined on the same substrate as polymer OLEDs<sup>2</sup> for a similar result. The big advantage of this approach is that an array of sensing elements can be combined with a display at very little incremental cost by simply adding “pixels that sense” in between the visible RGB display pixels. This would essentially augment a display with optical multitouch input “for free,” enabling truly widespread adoption of this exciting technology.

## Acknowledgments

We thank Stuart Taylor, Steve Bathiche, Andy Wilson, Turner Whitted and Otmar Hilliges for their invaluable input. 



## References

- Abileah, A., Green, P. Optical sensors embedded within AMLCD panel: design and applications. In *Proceedings of EDT'07* (San Diego, California, August 04, 2007), ACM, New York, NY.
- Bürgi, L. et al. Optical proximity and touch sensors based on monolithically integrated polymer photodiodes and polymer LEDs. *Org. Electron.* 7 (2006).
- Buxton, B. Multi-Touch Systems That I Have Known and Loved (2007), <http://www.billbuxton.com/multitouchOverview.html>.
- Dietz, P., Leigh, D. DiamondTouch: a multi-user touch technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology* (Orlando, FL, Nov. 11–14, 2001), UIST '01. ACM, NY, 219–226.
- Han, J.Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology* (Seattle, WA, Oct. 23–26, 2005), UIST '05. ACM, NY, 115–118.
- HP TouchSmart, <http://www.hp.com/united-states/campaigns/touchsmart/>.
- Hodges, S., Izadi, S., Butler, A., Rrustemi, A., Buxton, B. ThinSight: Versatile multi-touch sensing for thin form-factor displays. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, RI, Oct. 07–10, 2007), UIST '07. ACM, NY, 259–268.
- Hudson, S.E. 2004. Using light emitting diode arrays as touch-sensitive input and output devices. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (Santa Fe, NM, Oct. 24–27, 2004), UIST '04. ACM, NY, 287–290.
- Izadi, S. et al. C-Slate: A multi-touch and object recognition system for remote collaboration using horizontal surfaces. In *IEEE Workshop on Horizontal Interactive Human-Computer Systems* (Rhode Island, Oct. 2007), IEEE Tabletop 2007, IEEE, 3–10.
- Izadi, S. et al. Experiences with building a thin form-factor touch and tangible tabletop. In *IEEE Workshop on Horizontal Interactive Human-Computer Systems* (Amsterdam, Holland, Oct. 2008), Tabletop 2008, IEEE, 181–184.
- JazzMutant Lemur. [http://www.jazzmutant.com/lemur\\_overview.php](http://www.jazzmutant.com/lemur_overview.php).
- Jordà, S., Geiger, G., Alonso, M., Kaltenbrunner, M. The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (Baton Rouge, LA, Feb. 15–17, 2007), TEI '07. ACM, NY, 139–146.
- Krueger, M. *Artificial Reality 2*, Addison-Wesley Professional (1991). ISBN 0-201-52260-8
- Lee, S., Buxton, W., Smith, K.C. 5. A multi-touch three dimensional touch-sensitive tablet. *SIGCHI Bull.* 16, 4 (Apr. 1985), 21–25.
- van Loenen, E. et al. Entertable: a solution for social gaming experiences. In *Tangible Play Workshop, IUI Conference* (2007).
- Loviscach, J. Two-finger input with a standard touch screen. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, RI, USA, October 07–10, 2007), UIST '07. ACM, New York, NY, 169–172.
- Malik, S., Laszlo, J. Visual touchpad: A two-handed gestural input device. In *Proceedings of the 6th International Conference on Multimodal Interfaces* (State College, PA, Oct. 13–15, 2004), ICMI '04. ACM, NY, 289–296.
- Matsushita, N., Rekimoto, J. 1997. HoloWall: Designing a finger, hand, body, and object sensitive wall. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology* (Banff, Alberta, Canada, Oct. 14–17, 1997), UIST '97. ACM, NY, 209–210.
- Matsushita, N., Ayatsuka, Y., Rekimoto, J. 2000. Dual touch: A two-handed interface for pen-based PDAs. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology* (San Diego, California, US, Nov. 06–08, 2000), UIST '00. ACM, New York, NY, 211–212.
- N-trig, duo-touch sensor, <http://www.n-trig.com/>.
- Rekimoto, J. SmartSkin: An infrastructure for freehand manipulation on interactive surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves* (Minneapolis, MN, Apr. 20–25, 2002), CHI '02. ACM, NY, 113–120.
- Smart Technologies, <http://smarttech.com> (2009).
- Tactex Controls Inc. Array Sensors. [http://www.tactex.com/products\\_array.php](http://www.tactex.com/products_array.php).
- Matsushita, T. LCD with Finger Shadow Sensing, [http://www3.toshiba.co.jp/trm\\_dsp/press/2005/05-09-29.htm](http://www3.toshiba.co.jp/trm_dsp/press/2005/05-09-29.htm).
- Westerman, W. Hand Tracking, Finger Identification and Chordic Manipulation on a Multi-Touch Surface. PhD thesis, University of Delaware (1999).
- Wellner, P. Interacting with Paper on the Digital Desk. *CACM* 36, 7 (1993) 86–96.
- Wilson, A.D. TouchLight: An imaging touch screen and display for gesture-based interaction. In *Proceedings of the 6th International Conference on Multimodal Interfaces* (State College, PA, Oct. 13–15, 2004), ICMI '04. ACM, NY, 69–76.
- Wilson, A.D. PlayAnywhere: A compact interactive tabletop projection-vision system. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology* (Seattle, WA, Oct. 23–26, 2005), UIST '05. ACM, NY, 83–92.

Shahram Izadi, Steve Hodges, Alex Butler, Darren West, Alban Rrustemi, Mike Molloy, and William Buxton ([shahrami, shodges, dabj]@microsoft.com), Microsoft Research Cambridge, U.K.

© 2009 ACM 0001-0782/09/1200 \$10.00

# Take Advantage of ACM's Lifetime Membership Plan!

- ◆ **ACM Professional Members** can enjoy the convenience of making a single payment for their entire tenure as an ACM Member, and also be protected from future price increases by taking advantage of **ACM's Lifetime Membership** option.
- ◆ **ACM Lifetime Membership** dues may be tax deductible under certain circumstances, so becoming a Lifetime Member can have additional advantages if you act before the end of 2009. (Please consult with your tax advisor.)
- ◆ Lifetime Members receive a certificate of recognition suitable for framing, and enjoy all of the benefits of **ACM Professional Membership**.

Learn more and apply at:  
<http://www.acm.org/life>



Association for  
Computing Machinery

Advancing Computing as a Science & Profession