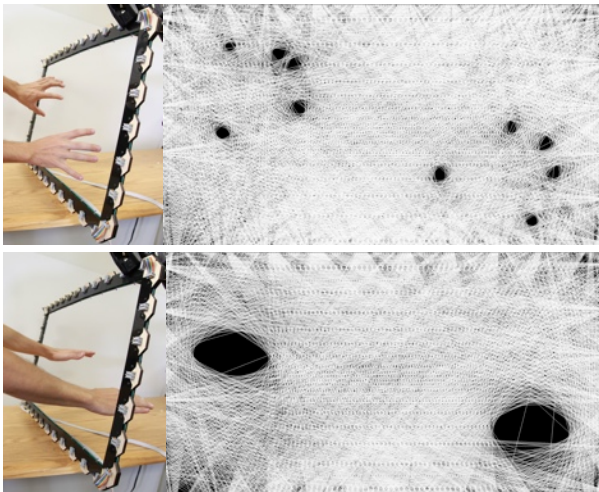# ZeroTouch: A Zero-Thickness Optical Multi-Touch Force Field

Figure 1

Zero-thickness visual hull sensing with ZeroTouch.

Jon Moeller
jmoeller@gmail.com

Andruid Kerne
andruid@ecologylab.net

Sashikanth Damaraju
damaraju@ecologylab.net

Interface Ecology Lab
Texas A&M University
Department of Computer Science
and Engineering

3112 TAMU
College Station, TX USA

Abstract

We present zero-thickness optical multi-touch sensing, a technique that simplifies sensor/display integration, and enables new forms of interaction not previously possible with other multi-touch sensing techniques.

Using low-cost modulated infrared sensors to quickly determine the visual hull of an interactive area, we enable robust real-time sensing of fingers and hands, even in the presence of strong ambient lighting. Our technology allows for 20+ fingers to be detected, many more than through prior visual hull techniques, and our use of wide-angle optoelectonics allows for excellent touch resolution, even in the corners of the sensor.

With the ability to track objects in free space, as well as its use as a traditional multi-touch sensor, ZeroTouch opens up a new world of interaction possibilities.

Keywords
Multi-touch, Sensor, Input Device, Visual Hull

ACM Classification Keywords
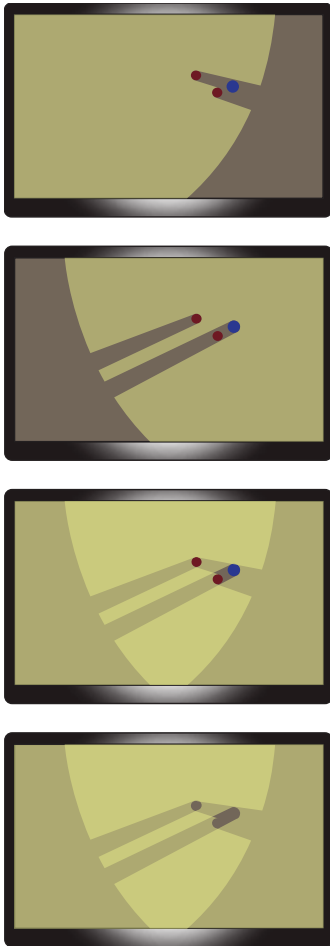H.5.2. User Interfaces: Input Devices and Strategies

General Terms
Design, Human Factors

Figure 2

Ambiguity of multiple touches in a corner camera style multi-touch monitor. The monitor can only "see" two touches, even though three are present. Because of the limited number of perspectives, only two fingers can be reliably detected due to occlusion problems.

Introduction

Multi-touch input has grown exponentially in popularity since the mass introduction of the technology in Apple's iPhone, iPad, and other multi-touch products. Multi-Touch integration with desktop systems has been significantly slower however, as most multi-touch technologies require hardware to be integrated within or behind the display, making them impossible to use with existing displays.

Capacitive technologies are limiting in that they require un-gloved human fingers to operate, and only work with specially designed styli [7]. Resistive screens are generally sensitive to any type of touch pressure, but suffer from display integration issues because of reduced light transmission through the touch surface due to materials used in their manufacture [7]. Optical systems range from bulky, camera-based systems to thinner optoelectronic approaches, but in general, all suffer from non-trivial display integration, with the exception of visual hull sensing techniques [2, 3, 6].

Recently commercialized optical multi-touch technologies, such as the HP Touchsmart [4], take a visual hull approach to sensing fingers and styli, using infrared cameras in the corners of a screen to track any objects touching the screen. However, this approach is limited due to touch-point ambiguity when using more than two fingers (see Figure 2). Most current flat-panel optical technologies can only distinguish between 3-4 fingers in the best case, and software support is usually limited to dual-touch.

Our technology solves this problem through the use of *point-to-point visual hull sensing*, which greatly increases the amount of information known about each touch point on the display, as well as providing simple integration with existing LCD displays.

Visual Hull Sensing

The visual hull of an object is the complete silhouette of an object, as seen from all sides. Corner-camera multi-touch sensors suffer from a lack of complete information about the visual hull of objects within the interaction area. In general, at least **n** viewpoints are needed to track **n** objects reliably. Figure 2 shows an example of incomplete information leading to ambiguity about specific touch points on a corner-camera style multi-touch screen.

While one touch point is correctly recognized, the other two touch points are incorrectly recognized as one continuous touch. This is because the correctly recognized finger is shadowing the space between the other two touch points, and is a result of the incomplete information offered by the corner-camera system.

To calculate a more complete visual hull, perspectives from all sides of the objects are needed. It is possible to increase the number of cameras to increase the number of perspectives on the screen, but this increase poses a problem: cameras are expensive; each additional camera added to the system significantly increases its cost.

Point-to-Point Visual Hull Sensing

To overcome these challenges, namely the need to gather visual hull information from a large number of perspectives while maintaining an appropriate cost structure, we use point-to-point visual hull sensing.

Figure 3

Disambiguation of multiple touches using point-to-point visual hull sensing. Each perspective only offers a limited amount of information, but when all perspectives are combined into a single image, touch points are easily distinguished.
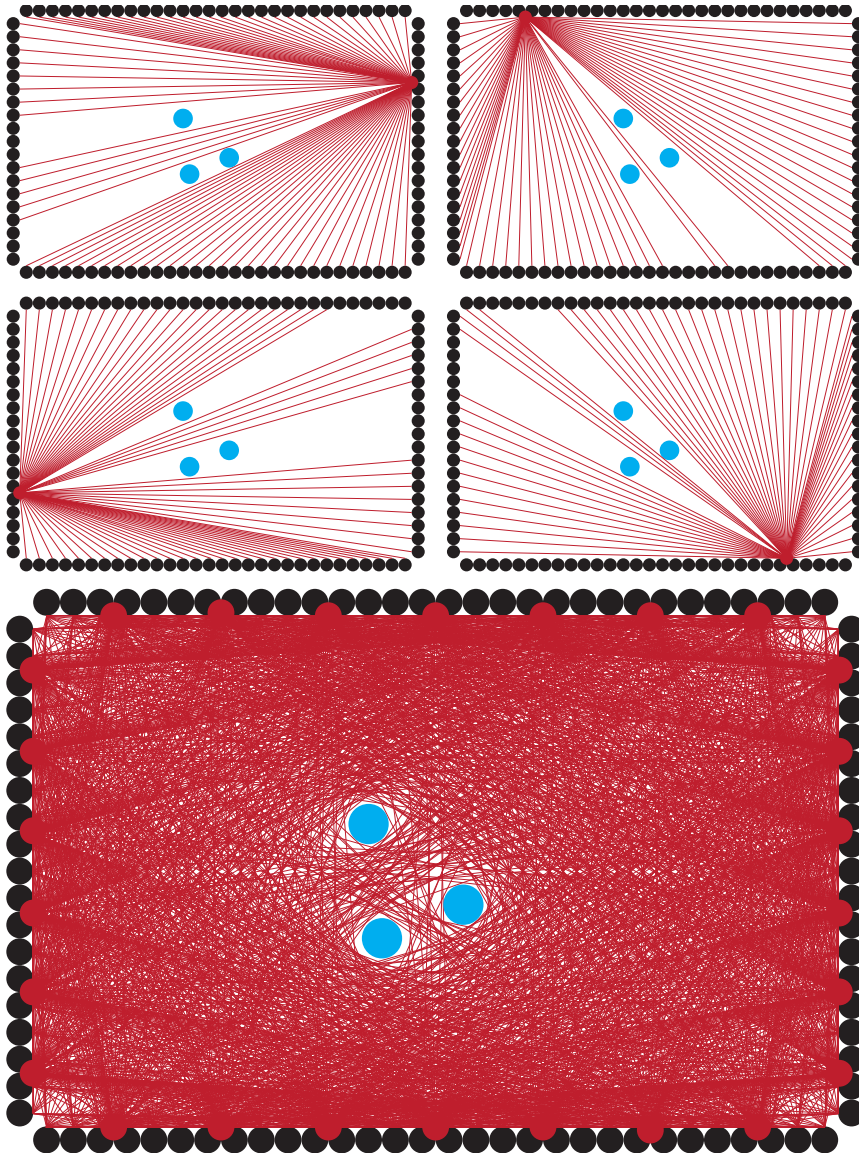
By using individual infrared sensors and LEDs, rather than multi-point receivers like cameras, we wrap the entire screen in a continuous sensor that provides more complete information about the visual hull of any objects within the interaction area. By surrounding the area with infrared sensors, and pulsing infrared LEDs at given positions along the sensor, a more complete visual hull of the interaction area is generated.

Figure 3 shows this principle at work, showing perspectives from four LEDs (top), along with the complete visual hull generated by the sum of all perspectives (bottom).

Each line crossing the screen is represented a binary 1 or 0, denoting the presence of absence of an object interrupting that line. In addition to giving a clear indication as to whether a light beam has been interrupted, this simplifies image processing and reduces the bandwidth required to send such data to the host PC.

Since the generated image is essentially a "picture" of the objects within the interaction area, traditional multi-touch image-processing techniques can be used to determine the location and size of touch points, such as those used in FTIR and other vision based multi-touch methods [1].

Sensor Technology

Using analog infrared sensors to implement point-to-point visual hull sensing is a possibility, but one must deal with the natural intensity variations that occur due to differences in the distances between individual sensors and sources. Aside from this, ambient light poses a big problem for these types of sensors, as they

Figure 4

Modular implementation of point-to-point visual hull sensing. Each module offers one additional perspective through the onboard LED, and 8 infrared sensors.

Modules are shown at actual size.

The prototype sensor, shown encircling the modules in this figure, consists of 32 modules in a daisy chain configuration. The sensor is approximately 28" diagonal.
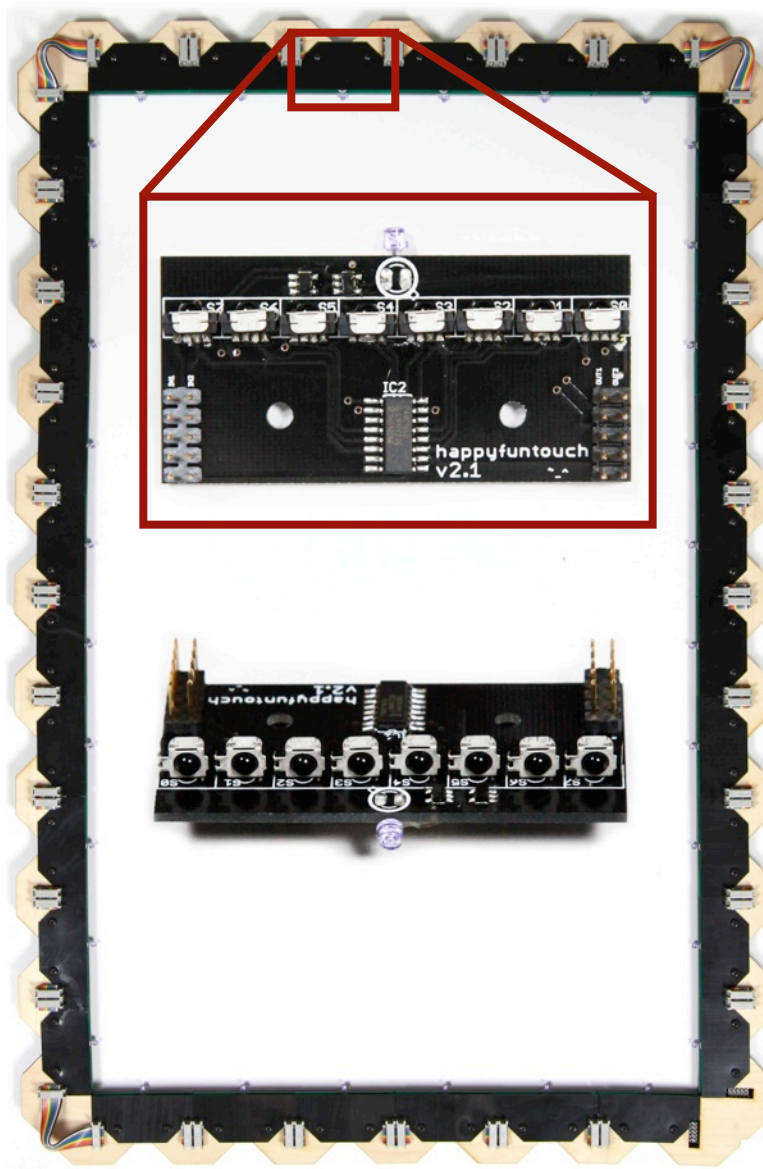
have no way of distinguishing ambient light variations from light variations coming from a real interaction.

To avoid these problems, we use commercially available modulated light sensors, typically used in television remote controls or garage door sensors. These sensors detect the presence or absence of light, but only if it is modulated at a specific frequency and for a specific amount of time. Using an internal band-pass filter and automatic gain control, they provide robust detection of signals, even in challenging ambient light conditions. In addition to this, the output from such a sensor is a binary 1 or 0, ideal for this application.

The other big advantage to IR remote sensors is that they can be read in parallel, allowing for very fast readout. While parallel readout is possible with traditional optoelectronics, parallel analog-to-digital conversion is much more expensive and much more data intensive than a simple binary readout. Each time an LED is pulsed at the appropriate frequency, a "snapshot" of the sensor is taken by simultaneously storing the values of all the sensors in a parallel-load shift register. This means that as more sensors are added to the screen (whether by increasing density or increasing size), The response time of the sensor remains essentially the same. Since the spatial resolution of the sensor is dominated by the sensors' spacing, and not the LEDs', there is no tradeoff between spatial resolution and response time.

Modular Design
Our prototype sensor is built with a number of individual modules (Figure 4). Each module contains 8 infrared remote sensors, and one infrared LED. Modules can be daisy-chained to create a full sensor of nearly

Figure 5

3-dimensional configuration of modules for 3D visual hull sensing.

any size. Our prototype sensor uses 32 modules, for a total of 32 perspectives and 256 individual sensors.

Modules can be arranged in nearly any configuration imaginable, allowing for both rectangular sensors for typical LCD displays as well as other odd shapes and combinations. Modules can also be arranged in 3-dimensional configurations, to allow for 3-dimensional point-to-point visual hull sensing, as shown in Figure 5.

Response Time

For a single perspective, from a single LED, the response time is the time it takes to pulse the LED and activate the sensor, plus the time it takes to transfer this data to the microcontroller. The shift registers operate in the multi-MHz frequency range, so the response time of the sensor is dominated by the pulse time for each LED.

IR receivers come in many flavors, each built to different specifications, depending on the application. They are manufactured with many different band-pass frequencies, the fastest in commercial production being 56kHz. The 56kHz sensors used in our prototypes require at least 6 pulses at this frequency for the sensor to activate, and about 10 cycles with no activity to deactivate. At 56kHz, this comes out to about 275μS required for each perspective.

In our prototype sensor, with 32 perspectives, a full update of all perspectives takes just under 10 ms.

Spatial Resolution

The spatial resolution of the sensor varies from point to point, since the effective grid resolution varies throughout the sensor. In general, resolution is better

in the center, and worse in the corners, because of the inherent density distribution of the light beams. That said, our screen has excellent corner resolution, around 1mm for single touches, and sub-millimeter accuracy for touches closer to the center. Multi-touch discrimination, the distance between two points before they can be recognized as such, is around 3 mm. The use of wide angle optoelectronics which enable light transmission between perpendicular sensors (as in the corners) allows for much better corner performance than in previous visual-hull techniques such as Scanning FTIR [6].

Multi-Touch Recognition & Performance

Data is transferred via USB to a host PC running Community Core Vision [1], which visualizes the data from the sensor by literally drawing lines between LEDs and activated sensors, and then applies standard image processing algorithms to determine blob position and
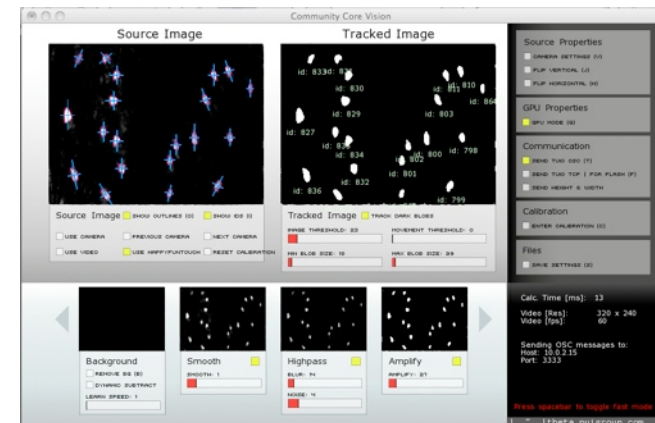
Figure 6

Community Core Vision screenshot showing 20 finger tracking with our prototype sensor.

size. These blob positions are output via the TUIO [5] protocol, and can be routed to a native Windows 7 multitouch driver, or a host of TUIO capable applications. Figure 6 shows a screenshot of Community Core Vision successfully tracking 20 fingers from four hands using our sensor.

Conclusion

Point-to-point visual hull sensing offers exciting new opportunities for multi-touch and gestural interfaces. ZeroTouch is a concrete embodiment of the point-to-point visual hull sensing principle, offering good spatial resolution, fast response time, and zero-touch activation. It works with both fingers and styli, and our 28" prototype can easily track 20+ objects at a time, more than enough for most use cases. In addition to operation as a traditional multi-touch screen, ZeroTouch can be used as an open-air interface, enabling new interaction techniques. Adding hover detection, for example, is simply a matter of adding an additional layer of sensors atop the base frame. We are also excited about the possibilities of 3-dimensional gestural interaction using point-to-point visual hull detection.

References
1. *Community Core Vision.* http://nuicode.com/ projects/tbeta
2. Hodges, S., Izadi, S., Butler, A., Rrustemi, A. and Buxton, B. 2007. ThinSight: versatile multi-touch sensing for thin form-factor displays. *Proc. UIST 2007*
3. Hofer, R., Naeff, D. and Kunz, A. 2009. FLATIR: FTIR multi-touch detection on a discrete distributed sensor array. *Proc. TEI 2009*
4. *HP Touchsmart.* http://www.hp.com/united-states/ campaigns/touchsmart/
5. Kaltenbrunner, M., Bovermann, T., Bencina, R. and Costanza, E. *TUIO - A Protocol for Table Based Tangible User Interfaces*. City, 2005.
6. Moeller, J. and Kerne, A. 2010. Scanning FTIR: unobtrusive optoelectronic multi-touch sensing through waveguide transmissivity imaging. *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*
7. Rosenberg, I. and Perlin, K. The UnMousePad: an interpolating multi-touch force-sensing input pad. *ACM Trans. Graph.*, 28, 3 2009), 1-9.