

ZeroTouch: An Optical Multi-Touch and Free-Air Interaction Architecture

Jon Moeller, Andruid Kerne

Interface Ecology Lab @ TAMU CSE

jmoeller@gmail.com, andruid@ecologylab.net

ABSTRACT

ZeroTouch (ZT) is a unique optical sensing technique and architecture that allows precision sensing of hands, fingers, and other objects within a constrained 2-dimensional plane. ZeroTouch provides tracking at 80 Hz, and up to 30 concurrent touch points. Integration with LCDs is trivial. While designed for multi-touch sensing, ZT enables other new modalities, such as pen+touch and free-air interaction. In this paper, we contextualize ZT innovations with a review of other flat-panel sensing technologies. We present the modular sensing architecture behind ZT, and examine early diverse uses of ZT sensing.

Author Keywords

Multi-touch; ZeroTouch; Free-Air; Interaction; Sensing

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Input Devices*

General Terms

Design, Experimentation

INTRODUCTION

Flat-panel multi-touch technologies are slowly but surely scaling to larger and larger screen-sizes. What was once predominantly the domain of bulky vision-based camera/projector multi-touch systems is now being rapidly encroached upon by more space-efficient technologies, albeit at higher relative cost/screen area.

Recent developments in large-area flat-panel multi-touch sensing have used optical technologies. Whether by incorporation of optical sensors in the display itself, or by surrounding a display with optical sensors and transmitters, optical technologies are able to scale to large screen sizes.

Capacitive sensing technologies have recently scaled to large displays, enabling high-precision multi-touch at a scale once the sole realm of vision-based sensing. However, capacitive requires factory integration, and thus

is unsuitable for integration with the large contingent of non-multi-touch displays already in the world.

There are a number of technologies that enable multi-touch interaction on non-interactive displays in the market today, and the majority of these employ optical-based touch sensing. Some use cameras and computer vision techniques, and some use optical sensors and emitters to detect touch.

In this paper, we detail ZeroTouch (ZT), a hardware/software architecture for multi-touch sensing [16]. ZeroTouch is a flat-panel optical multitouch technology using a linear array of modulated light receivers which surround the periphery of a display to detect touch. It is designed with a modular architecture. A complete sensor is built from a number of smaller sensing modules, allowing a full sensor to be built at any practical size.

First, we present an overview of flat-panel optical multi-touch techniques, and position ZeroTouch amidst the multi-touch sensing landscape. Next, we go deeper into the technology of ZT, describing its modular architecture, sensing technique, and temporal and spatial resolution characteristics of the sensor. Finally, we develop application areas of ZT with case studies. We wrap up with a discussion and implications for the technology.

OPTICAL FLAT-PANEL SENSING TECHNOLOGIES

There are a few techniques for optical flat-panel sensing, some which sense from the sides, and some which sense directly behind or within the display itself. We will discuss some issues common to all forms of optoelectronic sensing, develop an overview of prior optoelectronic techniques, and situate ZT among them.

Common Issues in Optoelectronic Sensing

Among the wide variety of techniques for optoelectronic touch sensing, most suffer from a few common problems which can interfere with a system's success. Ambient light sensitivity is perhaps the most important noise factor in optoelectronic multi-touch systems, followed by active light interference.

Ambient Light Issues

Almost all optoelectronic techniques suffer from ambient lighting interference of some kind or another. Systems that don't use modulated light in their operation are the most sensitive, as stray ambient light reduces signal/noise ratio, introduces spurious touch points, or in the worst case, precludes the system from working at all.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI '12, May 5–10, 2012, Austin, Texas, USA

Copyright 2012 ACM 978-1-4503-1015-4/12/05...\$10.00

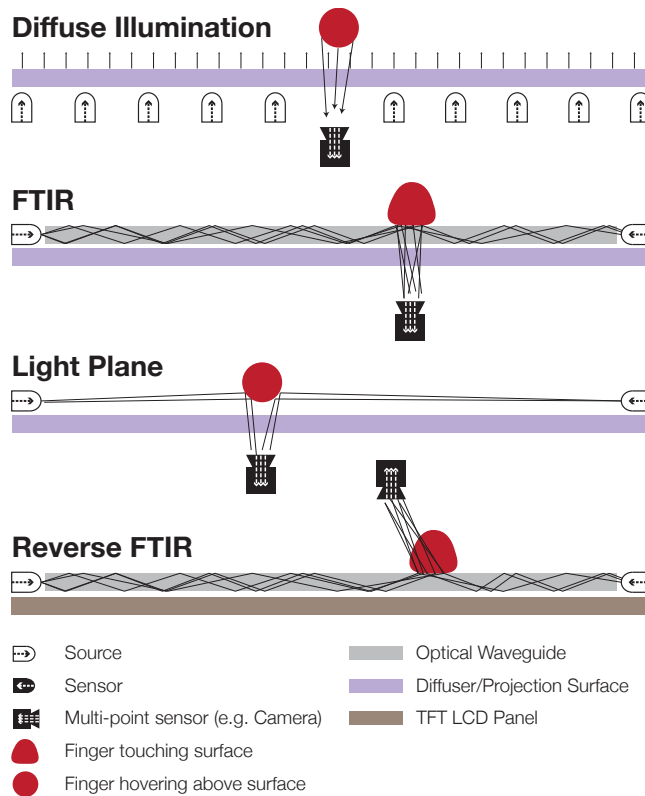


Figure 1: Vision-based techniques for multi-touch sensing.

One way to increase the effective signal/noise ratio of systems like this is to increase the effective illumination during the signal acquisition phase [3]. This form of light modulation still suffers from interference by high-intensity DC sources, like sunlight.

Systems using high-frequency modulated light, while well protected from DC sources such as tungsten lighting and sunlight, can still suffer ill effects from highly modulated fluorescent lights (like most CFLs), which are becoming increasingly popular as public requirements for minimizing energy consumption increase.

2D Vision-based Techniques

The most familiar instantiations of large multi-touch surfaces take the form of two-dimensional vision-based techniques. The commonality among these systems is use of a camera to directly image a 2-dimensional area for touches. The technique has been used in Perceptive Pixel's large-scale displays for CNN, Microsoft Surface, and many, many other home-brew and academic systems. Most, but not all vision-based systems require use of a projector and screen rather than a less bulky LCD display, so the camera can effectively image the interaction surface.

FTIR

Many FTIR systems use Han's 2004 method [6]. *Frustrated Total Internal Reflection* enables pressure-sensitive touch sensing by measuring the amount of light refracted outward from an infrared, edge-lit optical

waveguide. When a finger contacts the surface, the index of refraction changes; light escapes the waveguide, traveling toward the camera.

Inverted FTIR

A different take on FTIR, Inverted FTIR [4] uses the same imaging principle, but puts the camera in front of the display. While occlusion is a problem with this method, its advantage is that it can be used with a standard LCD panel.

Diffuse Illumination

Diffuse Illumination, the technology used in Microsoft Surface [13], uses a backlit diffuser to emit a fairly constant field of infrared light toward the interaction area. Reflected light from fingers or other objects is detected by a camera behind the screen.

Laser Light Plane

Laser light plane uses laser line emitters to reflect a thin plane of light into the interactive area. Light reflected off fingers or other objects is detected by a camera. Z-Touch [18] uses this method to simultaneously image 5 stacked planes of interaction by using a synchronized high-speed camera and five separate laser planes, enabling rudimentary depth sensing.

2D Sensor Integration Techniques

More recently, due to inherent drawbacks of camera and projector based methods for capturing multi-touch input, researchers have investigated two-dimensional arrays of sensors integrated with LCD displays. This integration allows for much smaller footprints, but requires major modification of an existing display, or fabrication of a display with entirely new technologies inside.

Thinsight

Microsoft Research's Thinsight [8] uses integrated LED and phototransistor pairs behind a traditional LCD display. While resolution is limited, it proves feasibility of sensing infrared light through an LCD.

FLATIR

FLATIR [9] uses sensors behind the screen, like Thinsight, but rather instead of putting LEDs there, they exploit FTIR principles, with an edge-lit acrylic sheet as the illumination source for sensors.

Pixelsense

Pixelsense [19], the technology of Microsoft Surface 2, is similar to Thinsight, except that the sensors are directly integrated into the LCD, rather than behind it. This offers a distinct advantage, in that infrared light only has to travel through the LC panel once, enabling higher-resolution imaging of the display area.

LED Array

Echtler's LED Array multi-touch display [5] eschews purpose-built infrared sensors, instead using LEDs themselves as the sensing elements. While the prototype

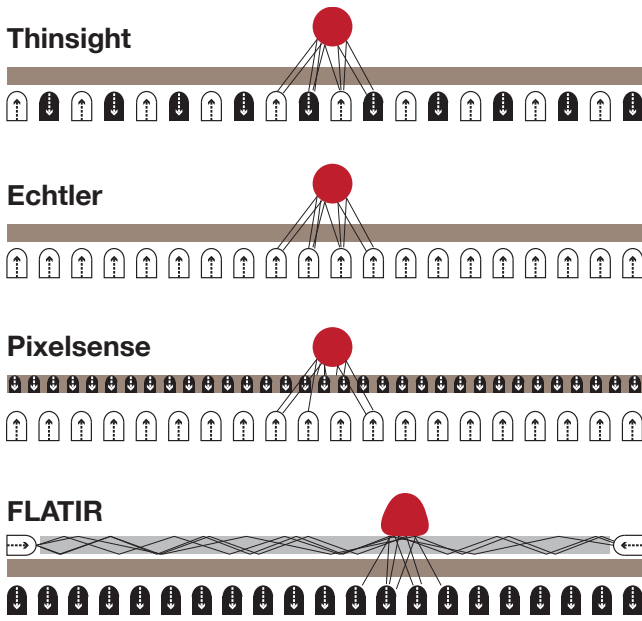


Figure 2: Illustration of sensor-integration techniques for multi-touch sensing.

only operated at ~10 Hertz, one of its big advantages is the use of modulated infrared light to reject ambient light.

Visual Hull Techniques

Visual hull techniques, of which ZeroTouch is an example, are the modern-day descendants of an old technique. Some of the first “touch” screens used similar techniques for single touch detection. By surrounding an area with infrared transmitters and receivers, determining the location of a touch is as simple as identifying which light beams are occluded in the x and y planes.

While this one-to-one arrangement of transmitters and receivers was once common, its limitations are obvious when more than one finger is on screen.

This problem is partially solved by establishing a many-to-many relationship between sensors and receivers, thereby generating a visual hull of the interactive area, as shown in Figure 3. While this partially solves the occlusion problem, inherent limitations remain in sensing concave objects, or objects within concavities. While two-dimensional imaging techniques don’t suffer from this problem, visual hull techniques are more than sufficient for touch sensing, as fingers are both minimally occluding and convex.

There are three main techniques for visual hull based multitouch sensing: corner camera, sensor occlusion, and medium interference (Figure 4).

Corner Camera

Corner camera systems position 2 or more infrared cameras in the corners of a display. Each camera has a different perspective on the interaction space; these perspectives are used to reconstruct a two-dimensional representation of the interaction area. In general, corner camera systems need one camera per touch-point to allow for ambiguity-free

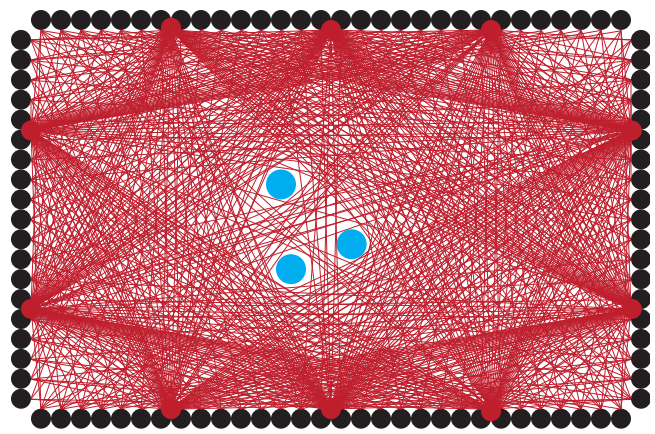


Figure 3: Many-to-Many visual Hull Sensing

sensing. Advanced algorithms can relax this constraint, but only to a certain extent. For reliable performance, it’s best to have at least one perspective per touch or object sensed.

Commercial systems using this approach include the Smart DViT [17], which uses four cameras to enable tracking of four simultaneous touches.

Sensor Occlusion

Sensor occlusion is a new spin on an old technology. A display is surrounded by a linear array of light sensors and sources. Each source is switched on one at a time, and all the sensors within range are read. Thus, each light source acts as an independent perspective, like a low-resolution corner camera.

As opposed to corner camera techniques, with a single constant light source, sensor occlusion techniques switch on only one, of a multitude of light sources, at a time. This necessitates time-slicing each individual perspective within a larger sampling window. For temporally coherent interaction at 60 Hz, this means sampling tens or hundreds of perspectives within a window of milliseconds, requiring precision timing and sensing techniques.

Philip’s Entertainable [11] was one of the first systems to apply this sensing technique to multi-touch, Jovanovic

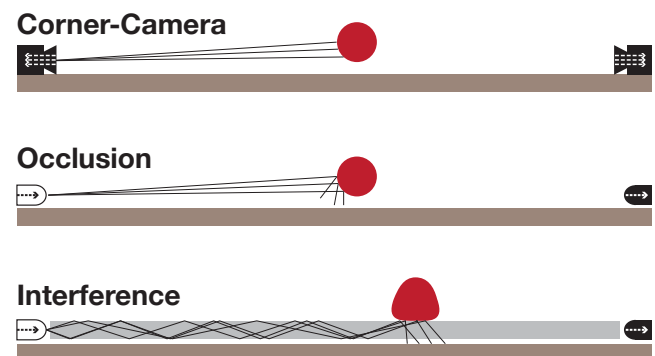


Figure 4: Illustration of visual-hull techniques for multi-touch sensing.

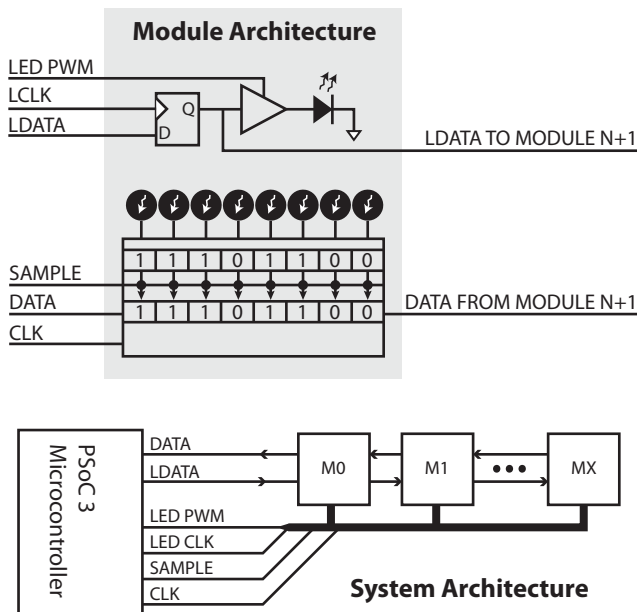


Figure 5: Module and System Architecture of ZeroTouch

investigated the geometric properties of sensors of this type [10,12].

Medium Interference

Medium interference techniques are operationally similar to sensor occlusion techniques. Rather than sensing the partial or complete occlusion of a light beam in mid-air, medium interference techniques analyze the reduction in light intensity as it travels through some medium. When a touch interferes with the medium, transmittance through the medium is decreased, as seen in Scanning FTIR [14]. This property makes it quite similar to computed tomography, since the intensity at a given point is the integral of the amount of light occluded, as opposed to a simple binary visual hull.

ZEROTOUCH

ZeroTouch is a sensor occlusion multi-touch sensing technique. The use of modulated infrared light sensors rather than discrete photodiodes or phototransistors differentiates it from other visual hull techniques, like Scanning FTIR or Entertaible.

Both of these technologies use a 1:1 ratio of emitters to receivers, as opposed to the 1:8 ratio used in ZeroTouch. Commercially available modulated light sensors take longer to respond than simple phototransistors, but can be sampled in parallel without expensive A/D conversion. By using a larger number of receivers than emitters, the ZeroTouch architecture allows for good temporal resolution, despite the slower response time of the receivers. In addition to this, the signal-noise ratio for binary occlusion data is much better when using modulated light receivers.

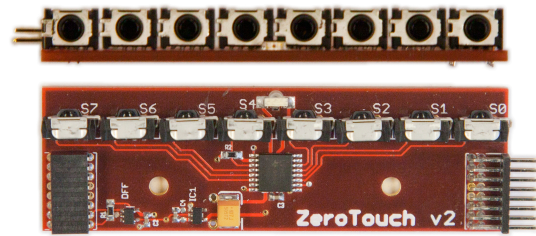


Figure 6: ZeroTouch Module (Actual Size)

This gives ZeroTouch several advantages in terms of sensor construction, electrical interconnect complexity, and ambient light rejection.

Modular Architecture

ZeroTouch develops a modular architecture in which many modules can be daisy-chained in most any configuration to construct a complete sensor. The chain of interconnected modules is, in turn, connected to a Cypress Programmable System on a Chip (PSoC) micro-controller. A diagram of the architecture is shown in Figure 5.

Each module contains eight modulated infrared light receivers, and a single infrared LED. A parallel load shift register simultaneously samples output from each receiver. A D flip-flop and buffer control an LED that acts as a distributed shift register when modules are chained.

The daisy-chain architecture allows sensors to be built as large as necessary using the same basic components. The 8-1 ratio between receivers and LEDs was chosen as a compromise between spatial and temporal resolution. For larger sensors, a 16-1 ratio may be more appropriate.

Module Design

Each ZeroTouch module is 2.425” long, with eight receivers spaced 0.3” apart. One LED sits in the middle of the board. Connectors on both sides allow power and data transfer to and from the board, enabling simple daisy-chaining of modules. The module can be seen at full scale in Figure 6.

Sensor Readout Cycle

Understanding the sensor readout cycle of ZeroTouch is necessary to discuss the temporal and spatial resolution tradeoffs that drove its current design.

Each cycle consists of reading data from every receiver, while pulsing one LED in the system. We define a scanline as the straight-line optical path from a single sensor to an LED. A single cycle generates scanline occlusion data from one LED’s perspective in the system (Figure 8). A frame of data is comprised of cycle data for each LED in the system. The operations completed in a single cycle are described below. Figure 7 shows a timing waveform diagram of the sensor readout cycle.

Step 1. Initialize LED Array

Each discrete flip-flop contains a bit determining whether or not to turn on an LED when its output is enabled. These flip-flops are loaded with all zeros, except for the one LED we are illuminating during this cycle.

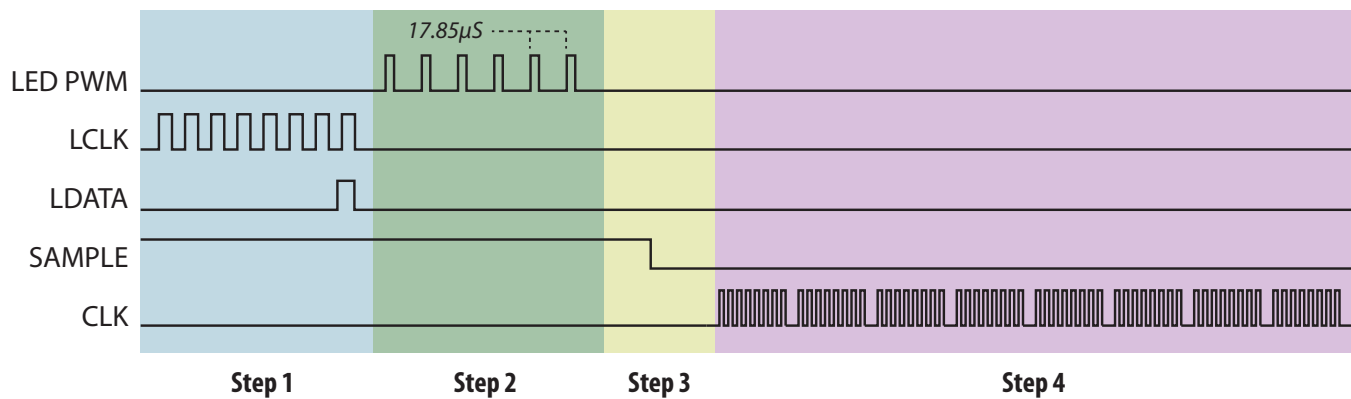


Figure 7: Timing Diagram for one cycle of an eight-module ZeroTouch Configuration. Timing is not to scale. In reality, Step 2 would take over 95% of the time for a single eight-module cycle, using a 20 MHz shift frequency.

Step 2. Pulse LED

Each LED in the chain is connected to a buffer with tri-state output. After the LED Array is initialized, the output enable pin of the buffer is pulsed at 56kHz, the center frequency of the modulated light sensors used in the system. This pulse happens for 6-8 56kHz cycles, depending on how the system is calibrated.

Step 3. Latch Receiver Data into Shift Registers

All receivers in the system are directly attached to a parallel-load shift register. When the latch line is driven low, the data at the pin of each receiver is loaded into its shift register. This effectively takes a snapshot of all receiver data outputs, from the given LED.

Step 4. Read Data from Shift Registers

Finally, all the data latched into the shift registers is read by the micro-controller, and placed into an internal RAM buffer. Because of tight timing constraints, data is read using a Verilog component programmed into one of the PSoC's digital blocks. The memory transfer takes place via high-priority DMA transactions.

After all the data is extracted from the receivers, the cycle repeats, this time initializing the LED array so that the next LED in sequence will be illuminated, and so on. Only one LED is activated at a time, so occlusion data is easily determined and localized. This is a key factor in the spatial and temporal resolution tradeoff in ZeroTouch.

Spatial/Temporal Resolution Tradeoffs

Because only one LED can be illuminated during an acquisition cycle, there is an inherent tradeoff between temporal and spatial resolution in a ZeroTouch system. The total frame capture time is dependent on the number of independent projections gathered by the system.

This means that while the number of receivers in a ZeroTouch system is theoretically unlimited, the number of independent LEDs that can be used to activate the receivers is limited, if temporal coherency is needed. That said, there is also a practical limit on the number of receivers used in the current daisy-chain architecture.

Sensor Clock Limits

After step 2 in the cycle, the receivers have a “cool down” period before they can be activated again to take a reading. This is a requisite of the internal signal conditioning circuitry inherent in each infrared receiver. This cool down period varies, depending on ambient light conditions and strength of the original pulses, but is generally 6-8 cycles of the 56kHz frequency of the receiver. This means that all data loaded into the shift registers must be shifted out and processed by the micro-controller during the $\sim 125\mu\text{s}$ cool down. While this is not an issue for ZeroTouch sensors of smaller size, it becomes significant as you scale the number of receivers. The PSoC can handle data rates of up to 40 MHz, so the largest practical sensor in a single daisy chain would contain 2048 receivers; this is more than enough for touch screen applications of nearly any size, and also reaches bounds of affordability in terms of hardware cost.

LED Perspective Limitations

In a ZeroTouch sensor, each LED is activated individually. All receivers are read simultaneously during this activation period. We call this activation and reading of a single LED's occlusions a single *perspective*.

So, while the cycle time is well defined for a ZeroTouch module, at around $275\mu\text{s}$ for activation, sensing, and data readout, the complete frame rate of the system depends on the total number of active perspectives used in the system. For example, in a 128 module system (approximately 108” diagonal) each perspective can be read in $275\mu\text{s}$. However, capturing data from 128 perspectives for a complete frame of data would take $\sim 35\text{ms}$, giving a frame rate of ~ 28 fps.

This can be ameliorated by using fewer active LEDs, enabling faster readout of large ZeroTouch sensors; this comes with a spatial resolution cost. Keeping the number of LEDs constant and increasing the number of infrared sensors enables larger ZeroTouch sensors to be constructed with equivalent temporal resolution.

A system with 512 infrared receivers, enough for a 55” display, can be read at 80 fps using 32 active LEDs. There are roughly 12,000 *scanlines* in a system of this size, compared to roughly 6,000 *scanlines* in our 27”, 256 sensor

prototype. While this means the resolution per area is halved, if there are an equal number of pixels on each display, for example, if both were 1080p or 1440p, the number of *scanlines* per pixel is doubled.

Because the receiver spacing remains the same regardless of screen size, the touch resolution (spacing between scanlines for a given perspective) remains nearly the same. While adding additional LEDs increases touch resolution to some extent, the main reason to do so is to increase the total number of fingers that can be sensed at a given time.

It's important to note that a detected touch is a linear combination of occluded scanlines. There are many more such combinations of scanlines than there are individual scanlines, so the resolution of touch tracking is much better than the number of scanlines divided by the inches it crosses.

Image Reconstruction and Touch Sensing

After data is acquired from the sensor and sent to the computer via USB, an image-based reconstruction algorithm processes the raw sensor data in order to use vision-based techniques for touchpoint recognition and tracking

Reconstruction Algorithm

The reconstruction algorithm is, in a sense, the inverse operation of the data acquisition stage. The touch image is reconstructed by visualizing the binary occlusion data gathered from the sensor.

The first step is generating a two-dimensional model of the ZeroTouch sensor that reflects the physical orientation of each module within the system. Given that most ZeroTouch sensors will be constructed in a rectangular form factor, there are convenience methods that can automatically generate the geometry for a rectangular frame, given the number of ZeroTouch modules used in the X and Y directions.

This two-dimensional model includes coordinates for each sensor and LED. The coordinate system is based around the ZeroTouch module, such that each module is 1 unit length, and so each sensor is 1/8th of a unit length.

After geometry is computed, and we know where all receivers and LEDs are located in our reconstruction plane, visualization of the *scanlines* can be as simple as drawing lines from each LED to each activated, non-occluded receiver. A more accurate approach involves drawing triangles from each LED to each non-occluded receiver. This is the approach taken in ZeroTouch, where each *scanline* is represented by a triangle with one vertex starting at an active LED, and the other two vertices at the edges of each sensor.

The visualized data from one LED can be seen in Figure 8, with one LED/sensor triangle shown in red. In this figure, note the imperfect reception pattern displayed by the receiver array. Not all receivers are activated by a single LED, even with no occluding touches. The receivers are self-calibrating, having internal automatic gain control, and

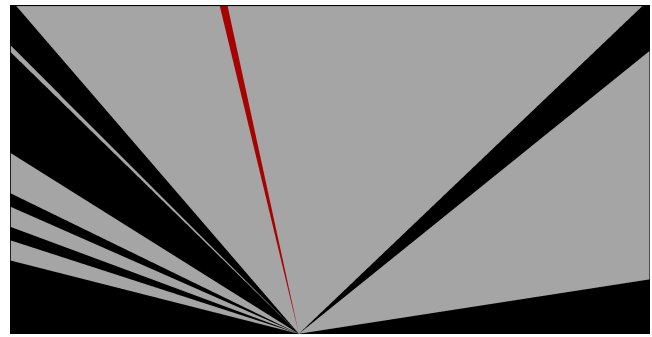


Figure 8: Visualized data from one LED. A single LED/receiver scanline is shown in red.

so some receivers are naturally more or less sensitive than others.

The maximum angle between an LED and receiver is 90 degrees in a rectangular ZeroTouch frame. Even at this 90 degree angle, receivers retain 30% of their nominal sensitivity. Coupled with wide-angle 160° LEDs, this ensures that light from an LED on one side of a ZeroTouch sensor will be seen by most receivers on the other three sides.

Ideally, all receivers should be activated by a single LED if there are no occlusions. However, the performance of the system is not compromised by small variations in reception performance, because each point on the display has many scanlines passing through it. For a touch to be recognized, all scanlines passing through a particular area must be occluded.

Image Processing and Touch Detection

After the raw sensor data has been visualized as a two-dimensional touch image, standard vision-based techniques for touchpoint recognition and tracking are used to determine the location, size, and movement of touches.

We use an enhanced version of Community Core Vision (CCV), an open-source vision-based multi-touch toolkit for touch recognition, originally designed for use with camera-based multi-touch systems. However, the algorithms for touch detection and tracking work equally well with ZeroTouch's reconstructed image.

After an image is created, it passes through a set of image filters to reduce noise and enhance tracking. First, the image is passed through a Gaussian smoothing filter. After this stage, the image passes through a thresholding filter to create a binary touch image, which is then used by the touch recognizer and tracker. Finally, the thresholded touch image is passed through a contour finding algorithm, which identifies touches in the image. Touches are passed to a blob tracker, which handles identification and state history for each touch.

Occlusion Issues

Occlusion can be a problem in ZeroTouch, because of the limitations of visual hull capture. Concavities in objects cannot be detected, and shadowing problems can be

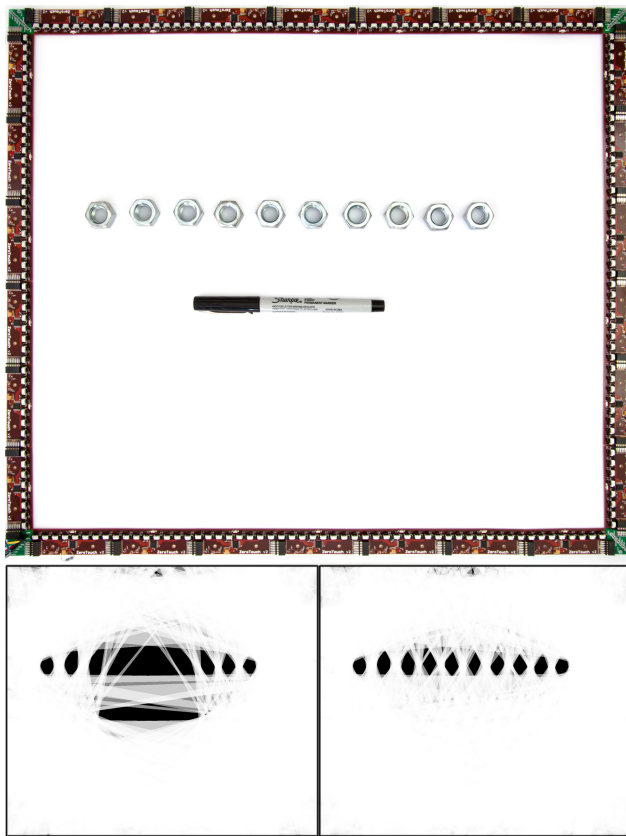


Figure 9: Example of occlusion problems that can happen with ZeroTouch. The data from the sensor shown above is visualized both with the pen (left) and without (right).

observed, as shown in Figure 9. However, because fingertips are small, convex objects, shadowing is typically a non-issue. One situation where it can become a problem is when a user's palm comes in contact with the screen or activates the sensor.

NEW MODALITIES FOR NUI INTERACTION

ZeroTouch can serve as an enabling component in various modalities of natural user interaction (NUI), either as a free-air interactive sensor or as an add-on multi-touch sensor. Here we present several NUI modalities enabled by ZeroTouch, each with design implications and insights.

Pen and Touch Interaction

The combination of pen and touch provides unique new opportunities for HCI [7]. We have successfully integrated ZeroTouch with a 21" Wacom Cintiq tablet display. ZT is calibrated such that the pen is not detected as a touch, so no disambiguation on the software side is necessary.

We demonstrated two pen + touch applications, Pen-in-Hand Command and Art Piles, at CHI 2011 Interactivity. Several hundred attendees interacted with both, experiencing ZT in real-world interaction scenarios. While this was not a controlled user study, their observations were an excellent indication of the utility and impact of the technology, especially given the experience of the average

attendee. We present initial interaction possibilities, user feedback, and observations.

Users at CHI were enthusiastic about the responsiveness of the system and its ease of integration with Wacom stylus displays. The combination of pen and touch interaction is a burgeoning topic in HCI. A number of researchers used the system extensively, and were interested in obtaining a ZeroTouch sensor specifically for this purpose.

Pen in Hand Command

Pen in Hand Command is a real-time strategy (RTS) game, based on Spring, an open-source RTS engine. We divide tasks in the game between macro-level tasks such as map navigation, and micro-level tasks such as unit manipulation and command invocation.

Macro-level tasks are performed using multi-touch interactions. Panning, zooming, and rotation are fully supported with single-touch pans and two-finger pinch to zoom and rotate. A third gesture, using three fingers is used for controlling camera tilt in the game.

The pen is used for fine-grained interactions, such as unit selection and manipulation. Lasso selection of units is supported, and a marking menu allows for in-context command invocation.

The pen can be used simultaneously with touch. One behavior observed in informal studies was users' fluid shifting between pen and touch interaction. Participants would hold the pen in their hand with one finger while simultaneously using the other fingers to perform macro-level operations. Bimanual use was also seen, dividing labor between pen in the dominant hand, and touch in the non-dominant.

ArtPiles

ArtPiles is another pen/touch interactive application, aimed at enabling museum curators to explore collections of items, form ideas about exhibits, and interact with metadata about the collections within a fluid interactive space. Pen gestures, such as drag and cross [1], were implemented for group formation. Combination pen/hand interactions were used for operations like grouping and pile formation.

Free-Air Interaction

We have used ZeroTouch in several *free-air* interaction [15] scenarios. ZeroTouch enables precise sensing of both fingers and hands in free air, although this sensing is constrained to the two-dimensional plane defined by the sensor.

intangibleCanvas

intangibleCanvas enables artists to paint on a remotely projected canvas by using hands, fingers, and other objects inside the ZeroTouch sensor [15]. By placing the sensor in direct line of sight with the display, visual connectedness with the remote canvas is established.

An iPhone application allows for color and brush selection, enabling expressive multi-modal finger painting at large scales. Ambiguous activation threshold issues were

overcome using a special targeting mode, activated by turning the iPhone palette face down. This mode allows the user to see where he/she is actually interacting with the sensor before activating the flow of paint by turning the phone display-side up.

intangibleCanvas was also demonstrated at CHI 2011 Interactivity. Attendees seemed enthusiastic about the possibilities of precision free-air interaction. Our most common complaint in this modality was the lack of feedback as to whether a user's hands were actually inside the sensor. This tracks well with our earlier observations during a more formal user study, where activation feedback was noted as a pressing need for this interaction modality [15].

Timepiece Explorer

Timepiece Explorer [20] is a gestural video browser created in collaboration with Zigelbaum+Coelho, for the international advertising agency Publicis.

Timepiece Explorer mounted the ZeroTouch sensor on a podium to track hands inside the interaction area. On the main screen, four video slices are shown. The user moves her hand around the space to select a video to play. The space responds kinetically to the user's movements, creating a fluid interaction space with affordances to aid the user in selecting a video for full screen playback.

In full-screen playback mode, the user can fast-forward and rewind video by manipulating an inertial playhead. The interaction area acts like a stream of water, playing video at normal speed with no interaction, and enabling fast-forward and reverse playback depending on the velocity and direction of user interactions inside the sensor.

Timepiece Explorer was used at a day-long event sponsored by Cartier: Fast Company's Most Creative People in Business 2011. It was the first real-world stress test of the system in practice. It ran the duration of the 8 hour event. Over 40 attendees interacted with the Timepiece Explorer installation during the event. Attendees who interacted with the system were generally impressed with the application, and with the tracking accuracy of the sensor.

Multi-Touch Augmentation

Besides adding multi-touch capabilities to an existing interactive display like a Wacom Tablet, ZeroTouch can be used to provide multi-touch input for haptic feedback devices, where the technology used in the device precludes other forms of flat-panel multi-touch sensing.

For example, TeslaTouch is a haptic feedback system for touch displays based on electrovibration [2]. However, because the system uses a technology similar to capacitive touch sensing to pass currents through the user's hands, integration of such a technique with capacitive multi-touch sensing could be challenging. In the system presented, they use rear projection and a vision-based multi-touch system to recognize touch. ZeroTouch might offer a compelling alternative, enabling the use of TeslaTouch, or other haptic

feedback technologies, with flat-panel LCDs and precision multi-touch sensing.

Hover Sensing

Another valuable use for ZeroTouch is the augmentation of existing multi-touch displays to enable hover sensing. While some vision-based systems allow for limited sensing of hover, it is most often a blurry view of what interactions are actually taking place above the surface.

ZeroTouch enables precision hover tracking when used with other multi-touch technologies, or when using two stacked ZeroTouch sensors. While stacking ZeroTouch sensors incurs a temporal or spatial resolution penalty because only one LED on each layer can be activated at a time, it may offer a compelling solution for a touch/hover based display.

Using ZeroTouch with another type of multi-touch sensing technology on the other hand offers the same precision tracking that ZeroTouch offers as a multi-touch system, but at a hover level. FTIR [6] is the most suitable technique for incorporation of hover, since there is little infrared light scattered toward the ZeroTouch's interactive area. Direct Illumination is also possible, although some tuning of the system or shielding may be required such that the active illumination doesn't interfere with ZeroTouch operation. However, this is only required if the lighting used for multi-touch sensing is modulated, as suggested in [3], because the infrared receivers used in ZeroTouch reject non-modulated light very well.

Other Uses

Because the ZeroTouch architecture is not limited solely to rectangular layouts, or even two-dimensional layouts, other potential interaction modalities are enabled by ZeroTouch.

Three Dimensional Sensing

Future work may investigate three-dimensional ZeroTouch interaction. By using ZeroTouch modules in non-planar configurations, three-dimensional visual-hull sensing is possible. This can be accomplished by stacking multiple planar layers, or by using non-planar arrangements in any configuration.

While three-dimensional sensing with ZeroTouch would not scale well to large interaction areas, especially compared to purpose-built vision-based solutions like Kinect, it has the potential to provide precision for smaller interaction areas.

Unique Form Factors

Because ZeroTouch modules can be arranged into in any number of form factors, possibilities abound. One interesting form factor is using two strips of ZeroTouch sensors to create a multi-touch or free-air interaction. There is no need to create a complete polygon of ZeroTouch sensors, as the one-to-many relationship between receivers and emitters enables multi-touch sensing even without enclosing the sides of the sensor, as shown in Figure 10. This configuration could have many uses such as hand

tracking in an interactive store shelf, or interaction with a heads-up display on an automobile windshield.

DISCUSSION

ZeroTouch has many applications in practice. However, it's important that we distill observations made while studying ZeroTouch in real-world scenarios, since the sensor's unique properties better enable some types of interactions than others. Free-air interaction with ZeroTouch allows for precise sensing, but currently doesn't provide enough feedback to take full advantage of its precision. However, using ZeroTouch as a multi-touch sensor is an ideal solution for making any existing display interactive, even when the display has other interaction modalities built-in.

Standalone Multi-Touch Use

First of all, ZeroTouch can be used to augment existing displays with multi-touch capability. In its current iteration, ZeroTouch tracks touches with pixel level accuracy on 27" 1440p displays. ZeroTouch enables 80 fps tracking at this accuracy with a touch latency of ~20 ms, when processing and tracking time is taken into consideration. No force is required to activate the sensor, and tracking is possible even with a gloved hand, enabling its use in potentially hazardous environments.

The infrared receivers we use in ZeroTouch have a range of up to 30 meters with sufficient illumination, enabling ZeroTouch sensors to be constructed for any practically sized displays.

The biggest advantages to using ZeroTouch as a multi-touch sensor is the ease of integration with existing LCD displays. The sensor is placed directly atop the display. No behind the screen components are necessary for its operation. There is no medium to interfere with the display.

Pen and Touch in Practice

Using ZeroTouch to augment stylus-capable displays such as Wacom's Cintiq Tablet is the easiest way to make a precision pen input display multi-touch capable. While ZeroTouch can track a stylus on its own, using a stylus device augmented with ZeroTouch enables pressure sensitive stylus input, easy disambiguation between stylus and touch events, and fast, accurate multi-touch tracking.

Without a dedicated stylus input device, disambiguation of touch and stylus events is a hard problem. With vision-based systems, disambiguation can be aided by using an active IR stylus. However, active IR styli lack the pressure sensitivity of dedicated stylus devices, such as those made by Wacom. Styli for capacitive touch screens are also available, but again, disambiguation of stylus and touch input presents a challenging problem that has so far gone unaddressed.

ZeroTouch offers a compelling solution to this problem because stylus and finger data are captured through entirely different channels, making disambiguation trivial.

While using purpose-built pen input devices presents an additional investment in a system like this, there are clearly advantages in synergistically combining a heterogeneity of



Figure 10: *Two-Sided Sensing with ZeroTouch. The data from the sensor shown above is visualized both with the pen (right) and without (left). Notice the distorted shape that the pen leaves, compared to Figure 9.*

sensing techniques, inasmuch as the techniques do not cause mutual interference.

Free-Air Interaction with ZeroTouch in Practice

In practice, free-air interaction with ZeroTouch has some inherent limitations due to the sensing technique used. Prior work with intangibleCanvas first assessed these limitations [15]. We further discuss them here.

The first big issue is activation feedback. Because ZeroTouch only senses interaction within a given two-dimensional plane, it is often not clear to the user when she has crossed into this plane of interaction. Visual feedback can help, but the absence of the tactile feedback of a touch surface is a challenge for the practical use of this modality.

This issue arose both in intangibleCanvas and in Timepiece Explorer. In intangibleCanvas, the issue was alleviated by the addition of a "targeting mode", activated by turning the iPhone color palette upside down. When this mode was enabled, the users interactions inside the sensor would not ink the canvas, but instead provide visual feedback of the user's interactions within the sensor. When the user turned the iPhone right-side-up to activate inking mode, ink would flow again at the point where the feedback was given in targeting mode.

One solution to this problem that we are beginning to explore is the user of laser-line modules to project a visible plane of laser light across the sensor to enable users to immediately know when they have crossed into the sensor's active area. Future work will investigate how this enhancement improves free-air interaction with ZeroTouch

as a compelling modality for applications like intangibleCanvas and Timepiece Explorer.

CONCLUSION

ZeroTouch enables entirely new forms of interaction and makes existing forms cheaper and more effective: both as a multi-touch sensor, and in other applications where precision sensing of hands and fingers within a two-dimensional plane is needed. It can be used for hover detection when coupled with an existing multi-touch display, or to augment a stylus-capable display with multi-touch interaction. Its form factor enables easy integration with existing LCD displays as a multi-touch sensor. The optical sensing technique also allows for integration with tactile and haptic displays that would otherwise be difficult to integrate with other multi-touch technologies.

ZeroTouch compares favorably in cost to vision-based solutions, especially at the screen sizes available in today's LCD panel market. At this point in time, ZT is much more affordable than capacitive-based multi-touch sensing on large screens. The tracking speed of ZeroTouch is faster than vision-based solutions, most of which operate at 60 fps.

ZT's utility in real-world multi-touch and free-air interaction has been validated by demonstration and use. Both expert and novice users alike have been impressed by the speed and accuracy of the sensor. This speaks to the usefulness of our technique in practice.

Architecturally, ZeroTouch offers a compelling solution for scalable touch sensing because the modules can be daisy chained, with little impact on temporal resolution, when the number of perspectives used in the system remains constant. The modular system also allows sensors to be configured in any number of form factors, contributing to the versatility of the system for any application where precise sensing of fingers and hands within is needed, and the interaction space is well defined.

ZeroTouch is an ideal architecture to support the use of large numbers of modulated infrared sensors, either in visual-hull capture or in other applications, because the modular architecture enables reconfigurability without increasing electrical interconnect complexity.

The synergetic combination of ZeroTouch with other input sensing modalities such as stylus and haptic feedback systems opens the door to new forms of natural user interaction. We are excited about these possibilities, and look forward to collaboration opportunities in this space.

ACKNOWLEDGEMENTS

Special thanks to Patrick Kane and Cypress Semiconductor for PSoC development kits and support. This work is supported by the NSF under grants IIS-074742 and IIS-0803854. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the NSF.

REFERENCES

1. Agarawala, A. and Balakrishnan, R. 2006. Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems (April 2006)*.
2. Bau, O. et al. 2010. TeslaTouch: electrovibration for touch surfaces. *UIST '10: Proceedings of the 23rd annual ACM symposium on User interface software and technology (October 2010)*.
3. Echtler, F. et al. 2009. A Short Guide to Modulated Light. *Proc. TEI 2009 (Cambridge, United Kingdom, 2009)*.
4. Echtler, F. et al. 2009. Inverted FTIR: Easy Multitouch Sensing for Flatscreens. *Proc. ITS 2009 (Banff, Canada, 2009)*.
5. Echtler, F. et al. 2010. An LED-based multitouch sensor for LCD screens. *Proc. TEI 2010 (Cambridge, Massachusetts, USA, 2010)*.
6. Han, J.Y. 2005. Low-cost multi-touch sensing through frustrated total internal reflection. *Proc. UIST 2005 (Seattle, WA, USA, 2005)*.
7. Hinckley, K. et al. 2010. Pen + Touch = New Tools. *Proc. UIST 2010 (New York, NY, USA, 2010)*.
8. Hodges, S. et al. 2007. ThinSight: versatile multi-touch sensing for thin form-factor displays. *Proc. UIST 2007 (Newport, Rhode Island, USA, 2007)*.
9. Hofer, R. et al. 2009. FLATIR: FTIR multi-touch detection on a discrete distributed sensor array. *Proc. TEI 2009 (Cambridge, United Kingdom, 2009)*.
10. Jovanovic, N. and Korst, J. 2008. Minimum blocking sets of circles for a set of lines in the plane. *Proc. CCCG 2008*.
11. Jovanovic, N. et al. 2009. Object Detection in Flatland. *ADVCOMP '09 (2009), 95–100*.
12. Jovanovic, N. et al. 2009. Candle in the woods: asymptotic bounds on minimum blocking sets. *SCG '09: Proceedings of the 25th annual symposium on Computational geometry (June 2009)*.
13. Microsoft Surface. <http://www.surface.com/>.
14. Moeller, J. and Kerne, A. 2010. Scanning FTIR: unobtrusive optoelectronic multi-touch sensing through waveguide transmissivity imaging. *Proc. TEI 2010 (Cambridge, Massachusetts, USA, 2010)*.
15. Moeller, J. et al. 2011. intangibleCanvas: free-air finger painting on a projected canvas. *CHI EA '11: Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (May 2011)*.
16. Moeller, J. et al. 2011. ZeroTouch: a zero-thickness optical multi-touch force field. *CHI EA '11: Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (May 2011)*.
17. Smart Technologies DVIT. <http://www.smart-technologies.com/dvit.html>.
18. Takeoka, Y. and Miyaki, T. 2010. Z-touch: an infrastructure for 3d gesture interaction in the proximity of tabletop surfaces. *Proc. ITS 2010 (2010)*.
19. The Power of PixelSense™. <http://www.microsoft.com/surface/en/us/pixelsense.aspx>.
20. ZIGELBAUM + COELHO - Timepiece Explorer. <http://zigelbaumcoelho.com/timepiece-explorer/>.