

Multi-Touch Surfaces: A Technical Guide

Technical Report TUM-I0833

Johannes Schöning*, Peter Brandl, Florian Daiber, Florian Echtler, Otmar Hilliges, Jonathan Hook, Markus Löchtefeld, Nima Motamedi, Laurence Muller, Patrick Olivier, Tim Roth, Ulrich von Zadow

ABSTRACT

Multi-touch interaction with computationally enhanced surfaces has received considerable recent attention. Approaches to the implementation of multi-touch interaction such as Frustrated Total Internal Reflection (FTIR) and Diffused Illumination (DI) have allowed for the low cost development of such surfaces, leading to a number of technology and application innovations. Although many of these techniques have been presented in an academic setting, the practicalities of building a high quality multi-touch enabled surface, both in terms of the software and hardware, are not trivial.

This document aims to summarize the knowledge and experience of developers of multi-touch technology who gathered at the Bootcamp on Construction & Implementation of Optical Multi-touch Surfaces at Tabletop 2008 in Amsterdam, and seeks to provide hints and practical advice to people seeking to “**build your own**” multi-touch surface. We mostly focus on technical aspects that are important in the construction of optical multi-touch surfaces, including: infrared illumination, silicone compliant surfaces, projection screens, cameras, filters, and projectors. In addition, we outline how to integrate this hardware to allow users to create a solid multi-touch surface, and provide an overview of existing software libraries for the implementation of multi-touch applications. In addition, we discuss the problem of latency introduced by the different parts of the system. A brief description of most of the common technologies to realize (multi-) touch surfaces is provided; however, the main focus is on those that utilise optical approaches.

Categories and Subject Descriptors

B.0 [Hardware General]: Multi-Touch Surfaces

Keywords

Multi-touch surfaces, Hardware for multi-touch surfaces, Software libraries for multi-touch surfaces, Multi-touch bootcamp in conjunction with IEEE Tabletops 2008

* Contact Address:

Johannes Schöning, Institute for Geoinformatics University of Münster, Weseler Str. 253, 48151 Münster, Germany, j.schoening@uni-muenster.de

Technical Report TUM-I0833



Figure 1: The joy of multi-touch interaction.

1. INTRODUCTION

Multi-touch technology is not entirely new, been available in different forms since the 1970s. Multiple patents [9, 23, 24, 28, 42] demonstrate how camera/sensor based touch sensitive surfaces can be constructed. Bill Buxton’s multi-touch webpage [3] gives a thorough overview of the underlying technologies as well as the history of multi-touch surfaces and interaction. It also provides a list of “traps” that people starting out with multi-touch interaction should be aware of. Of particular interest to Buxton are the possibilities afforded by today’s technology to exploit bi-manual interaction [4, 17] in a range of different application settings.

The rediscovery of the Frustrated Total Internal Reflection (FTIR) principle [13] has greatly accelerated the development of new multi-touch applications. In 2005 Han [13, 14] presented his low cost camera-based multi-touch sensing technique. Han’s YouTube demonstration captured the imagination of experts and laymen alike. His system was both cheap and easy to build, and illustrated a range of creatively applied multi-touch interaction techniques. In 2007 Apple presented their new mobile phone, the iPhone [1]. Where other touch based cellular phones only allow single point interaction, the iPhone used multi-touch technology. The resulting interaction techniques and interfaces have received considerable media attention and brought multi-

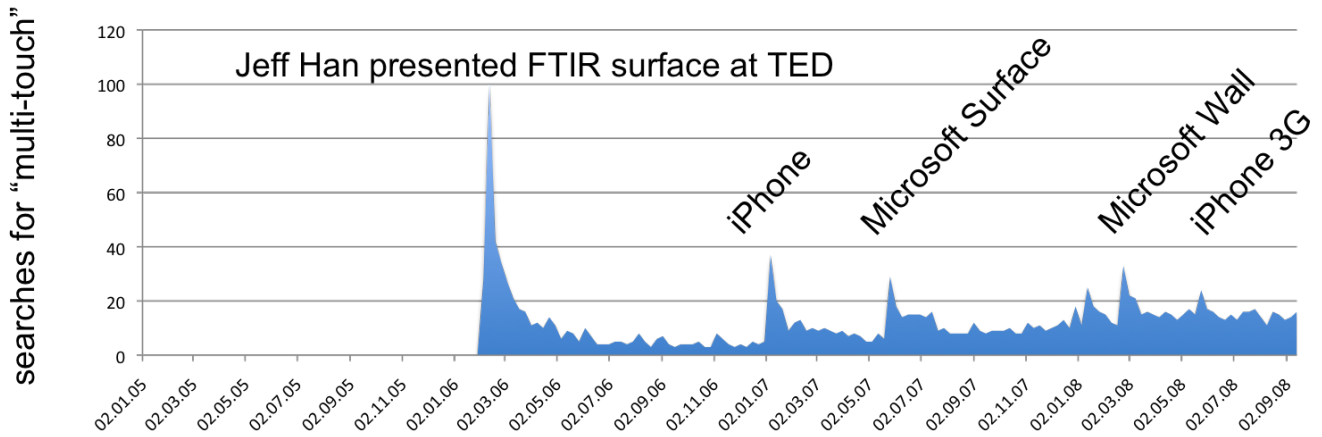


Figure 2: The popularity of the search terms “multi-touch” from the beginning of 2004 to now analysed with Google Trends and with different main events labelled.

touch interaction to the consumer electronics market.

Later in 2007 Microsoft presented their own version of a multi-touch table, MS Surface [26]. The table has the appearance of a coffee table with an interactive surface.

The sensing technique used in MS Surface is similar to the HoloWall [25] exploiting a diffuser which is attached to the screen material. The table surface is illuminated from behind with infrared light and when a user touches the table reflected infrared light is captured by cameras inside the table. Because of the use of multiple cameras, the input resolution is high enough to detect objects. It is interesting that this development can be depicted by using Google Trends [12, 44] or Google Insight Search [11] as can be seen in figure 2.

Despite these innovations many open questions for researchers remain: What are the benefits of multi-touch systems over single-touch systems? What are suitable applications? What kinds of applications are adequate for multi-touch systems? Are there more than interaction possibilities than “just” rotating and scaling photos or zooming into maps? Rather than addressing such questions, we focus on technical realities of building of optical multi-touch surfaces.

This document should therefore be read as a beginner’s guide and a pointer to the relevant literature.

The remainder of this paper is structured as follows: We start with a brief overview of existing (multi-)touch technologies. Section 3 focuses on the technical challenges users face when constructing a “**build your own**” multi-touch surface; namely (as described in section 3.1) infrared illumination, camera set-ups, filters, projectors, silicone compliant surfaces, projection screens and the integration of all this hardware into the final multi-touch surface. Section 3.2 focuses on existing software libraries which allow deployment of multi-touch applications upon optical multi-touch surfaces. Finally, Section 5 describes a selection of interesting projects currently utilising the technologies that we describe.

2. TOUCH TECHNOLOGIES

Before describing FTIR and DI there are a number of alternative technologies that can be used to construct multi-touch surfaces:

- Resistance Based Touch Surfaces
- Capacitance Based Touch Surfaces
- Surface Wave Touch Surfaces(SAW)

Unfortunately these technologies require industrial quality fabrication facilities to construct and are therefore not suitable for self-built surfaces, but are discussed for completeness.

2.1 Resistance Based Touch Surfaces

Resistance based touch surfaces generally consist of two conductive layers which are coated with substances such as indium tin oxide [45] These layers are separated by an insulating layer, usually made of tiny silicon dots (see figure 3 according [37]). The front of the panel is typically made of a flexible hard coated outer membrane while the back panel is often a glass substrate. A controller alternates between the layers, driving one with a specific (electric) current and measuring the current of the other. When users touch the display, the conductive layers are connected, establishing an electric current that is measured once horizontally and vertically by the controller in order to determine the exact position of a touch. Such touch surfaces have the advantage of low power consumption and are used in mobile devices such as the Nintendo DS [30], PDAs and digital cameras and can be operated both with fingers or a stylus.

However, resistance based surfaces provide a low clarity interactive surface (about 75%–85%) and additional screen protection cannot be applied without impacting on their functionality. More detailed information about classical resistance based (multi-) touch surfaces can be found in [8].

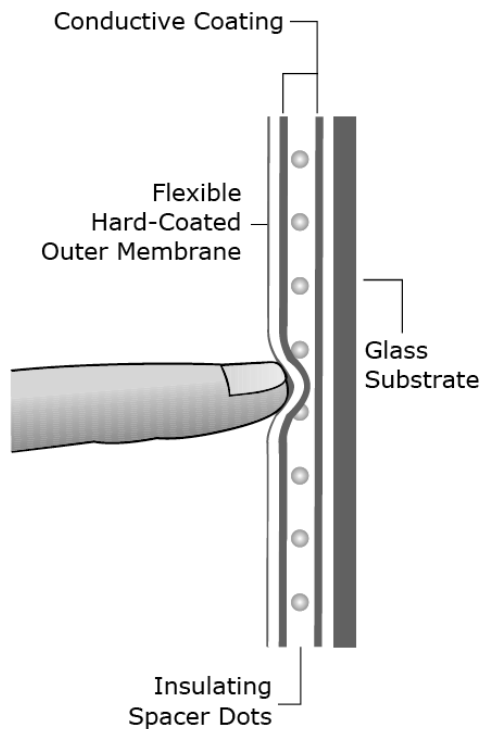


Figure 3: Schematic construction of a touch screen based on resistive technology according [37].

2.2 Capacitance Based Touch Surfaces

In general capacitance based (multi-) touch surfaces can be subdivided into two classes:

- Surface Capacitance
- Projected Capacitance

Both techniques were primarily developed for single touch interaction. One advantage of capacitive touch surfaces in comparison to other technologies is their high clarity; making capacitive touch surfaces very suitable for use in many kinds of touch displays beyond simple touch pads. Capacitive touch screens can also be operated by any conductive device and are hence not limited to finger based interaction. However, capacitive touch panels are relatively expensive to produce although they exhibit high durability and reliability. Consequently, capacitive based systems are often preferred for use in rough environments such as public displays and industrial applications.

It is possible to use such systems for multi-touch surfaces, but typically the number of simultaneous touches is limited by firmware or by the design of the controller. Finally, accuracy decreases when performing touches with more than one object. Having noted these general limitations, capacitance based technologies developed e.g. by MERL overcame many of these restrictions in order to allow many simultaneous touches; these are briefly described below.

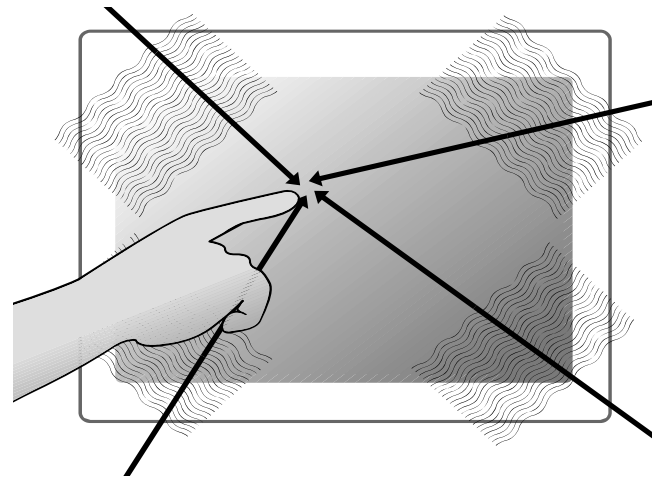


Figure 4: Surface Capacitive (Multi-) Touch Surfaces: Electrodes around the edges distribute voltage across the conductive layer creating an electric field. Touching the panel results in current drawn from each corner which is measured to define the position according [37].

2.2.1 Surface Capacitive Touch Surfaces

Surface capacitive touch panels consist of a uniform conductive coating on a glass layer. Compared to resistive technologies, a much higher clarity can be achieved by using indium tin oxide [45] as the conducting material (it is transparent as well as colourless when used in very thin layers). From each side of the touch panel electrodes maintain a precisely controlled store of electrons in the horizontal and vertical directions which sets up a uniform electric field across the conductive layer. As human fingers (or other conductive objects) are also electrical devices capable of storing charge and exhibiting electric fields, touching the panel results in a small transport of charge from the electric field of the panel to the field of the touching object. Current is drawn from each corner of the panel; this process is measured with sensors located in the corners, and a microprocessor interpolates an exact position of the touch based on the values measured (see figure 4). Panels based on surface capacitive technology can provide a high positional accuracy.

2.2.2 Projected Capacitive Touch Surfaces

Of the technologies we describe projected capacitive touch devices are the most expensive to produce. Their performance is rather worse than many of the other approaches we describe; however they afford superb mechanical resilience. Projected capacitive surfaces can also be covered by a non-conductive material (with a maximum thickness of around 20mm) without negatively impacting on their functionality. When used for (multi-) touch displays, as described by Rekimoto [32] a very thin grid of microphone wires is installed between two protective glass layers (see figure 6). When touched, capacitance forms between the finger and the sensor grid and the touch location can be computed based on the measured electrical characteristics of the grid layer. The accuracy of projected capacitive technology is similar to surface capacitive technology although light transmission is superior because the wire grid can be constructed such that it

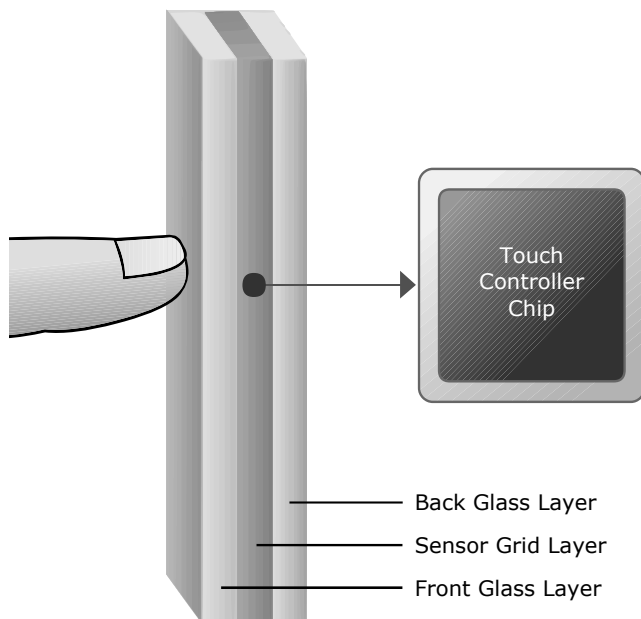


Figure 5: A thin grid layer is protected by two glass layers. Capacitance forms between the finger and the grid during a touch. The change of electrical properties is measured to determine the touching position according [37].

is nearly transparent. The technology is also highly suitable for rugged environments such as public installations, as a protective layer (such as thick glass) may be added without drastically decreasing the sensitivity. Finally, multiple simultaneous touches can be more easily interpreted compared to surface capacitive based technology.

MERL Capacitive Diamond Touch. In 2003 *DiamondTouch* was developed in the Mitsubishi Electric Research Laboratories (MERL). *DiamondTouch* was designed to support multiple touches, be tolerant of objects placed on the table, durable, un-encumbering, and inexpensive to manufacture [7]. The system has a number of distinctive characteristics:

- the ability to handle many touch points and users (only limited by the size of the table and the available space around it);
- the ability to identify which users are interacting with the surface;
- it is affected by debris objects like cups placed upon its surface
- it does not require additional devices for interaction (such as special pens).

Furthermore, the isolating layer between the antenna array and the user can be manufactured from a wide array of materials; hence it can be made to be very robust. For example, when a special fibre glass laminate is used, alcohol may be ignited on the surface without causing damage.

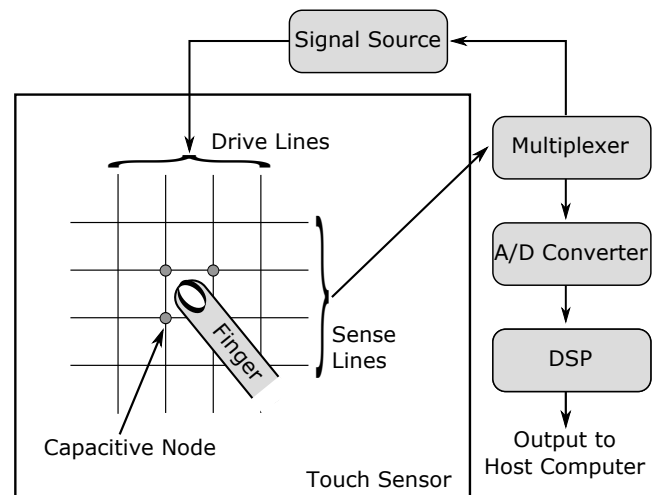


Figure 6: A simplified view of projected-capacitive touch screen adapted from Rekimoto.

Using capacitive coupling *DiamondTouch* is composed of a table with integrated antennas transmitting unique signals, a ceiling-mounted projector presents a display onto the table, one conductive chair connected with a receiver for each user and a computer. “When a user touches the table, a capacitively coupled circuit is completed. The circuit runs from the transmitter, through the touch point on the table surface, and through the user to the user’s receiver back to the transmitter.” [7]. *DiamondTouch* works by transmitting signals through antennas in the table; these signals are used to identify the parts of the table each user is touching. This information can then be used to calculate [7] the finger’s position. Usually a user touches several antennas at once. For this reason the signals have to be separable (in technical terms orthogonal). This can be achieved by frequency-division multiplexing, time-division multiplexing or code-division multiplexing. The antenna pattern consists of two layers similar in design, but with one rotated by ninety degrees. The rows/columns (antennas) of each layer are composed of diamond shapes connected in one direction and isolated in the other. In this way, the covered surface is maximised and the shielding effect minimised. Usually there is an antenna every five millimetres (which is the resulting minimum pointing accuracy). Due to image projection from above the only obstructions are shadows cast on the table by objects (i.e. hands or arms) inserted into the projector’s light beam.

2.3 Surface Wave Touch Surfaces (SAW)

Systems that use surface wave technology are similar to those that use infrared grid technology. Transmitting and receiving piezoelectric transducers, for both the X- and Y-axes, are mounted on a faceplate and ultra-sonic waves on a glass surface are created and directed by reflectors. By processing these to electronic signals and observing the changes when the faceplate is touched, it is possible to calculate the position of that interaction. Most SAW systems can support dual-touch.

2.4 Optical Based Touch Surfaces

Both optical and camera based-approaches share the same concept of processing and filtering captured images on patterns. As already discussed a number of systems are based on infrared illumination and as a result can suffer interference from ambient light in the environment. Due to their simple configuration optical approaches have the potential to be very robust.

2.4.1 Frustrated Total Internal Reflection (FTIR)

The rediscovery of the Frustrated Total Internal Reflection (FTIR) principle by Han [13, 14] in 2005 can be seen as a starting point for optical multi-touch systems. FTIR technology is based on optical total internal reflection within an interactive surface. Electromagnetic waves transmitted within an inner material are completely reflected at its boundary if both the inner material has a higher refractive index than the outer material and the angle of incidence at the boundary between the materials is small enough. Common FTIR set-ups have a transparent acrylic pane with a frame of LEDs around the side injecting infrared light. When the user touches the acrylic, the light escapes and is reflected at the finger's point of contact due to its higher refractive index; an infrared-sensitive camera at the back of the pane can clearly see these reflections. A basic set of computer vision algorithms is applied to the camera image to determine the location of the contact point. As the acrylic is transparent a projector can be located behind the surface (near to the camera) yielding a back-projected touch sensitive display. The general set-up of a FTIR system is illustrated in figure 7.

2.4.2 Diffuse Illumination (DI)

The hardware for Diffuse Illumination (DI) systems is similar to that for FTIR. Both techniques place a projector and infrared sensitive camera behind a projection surface. However, for DI, the infrared lighting is also placed behind the projection surface. This causes the area in front of the surface to be brightly lit in the infrared spectrum. Consequently, the camera picks up all objects in this area by their reflection of infrared light. This includes objects in proximity to as well as objects touching the surface. Touch detection exploits the fact that the projection surface diffuses light, blurring the images of objects at a distance. In contrast to FTIR, DI allows tracking and identification of objects as well as fingers. Objects can be identified using their shape or fiducials [5] (easily recognizable markers) printed on their bottom surfaces. Furthermore, any transparent surface (such as safety glass) can be placed between the projection screen and the user since the sensing does not rely on surface contact.

2.4.3 Diffused Surface Illumination (DSI)

Diffused Surface Illumination (DSI) addresses the problem of getting an even distribution of infrared light across the screen surface. A task which is typically achieved in DI set-ups by using a small number of (two or three) Infrared Illuminators. In DSI Tim Roth [33] proposed the use of a special acrylic that incorporates small particles that act as tiny mirrors. So when IR light shines into the edges of this material it is redirected and evenly spread across the surface. Figure 9 illustrates the DSI set-up. In his blog Roth summarizes the pro and cons of this approach:

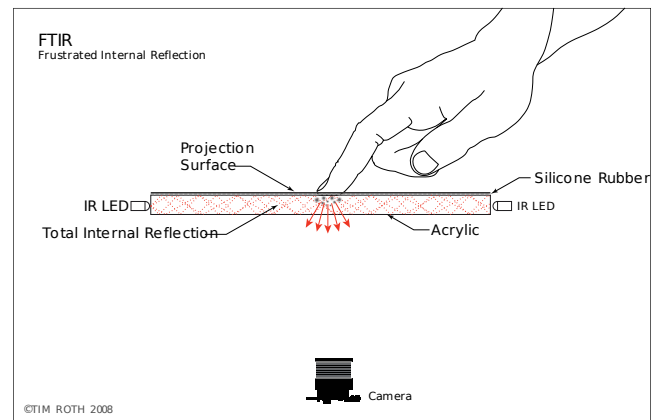


Figure 7: General set-up of a FTIR system

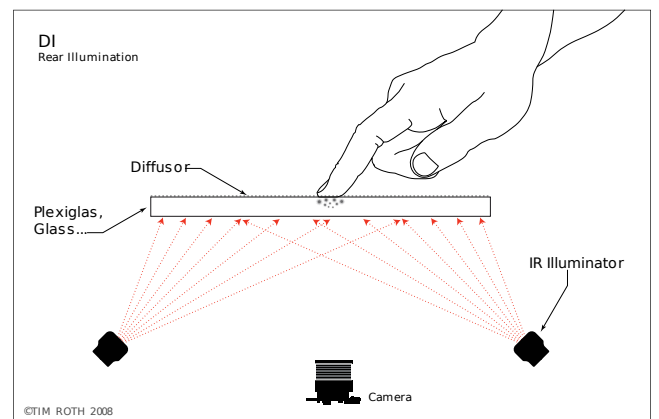


Figure 8: General set-up of a Diffuse Illumination system

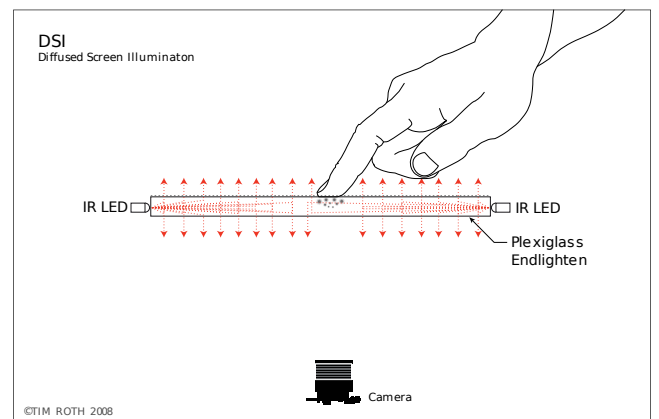


Figure 9: General set-up of a Diffused Surface Illumination system

- Advantages of DSI**
1. no problems getting an even distribution of light, easy set-up;
 2. an FTIR set-up can be converted to DI easily;
 3. no need to build a special case for the set-up (like other DI set-ups);
 4. fiducial tracking is possible.

- Disadvantages of DSI**
1. less contrast compared to normal DI set-ups as the surface material also redirects the IR towards the camera;
 2. potentially more problems with ambient IR because of less contrast;
 3. possible size restrictions because of the softness of the surface material.

2.4.4 LCD Enabled MT Surfaces with Optical-Based Sensors

LCD monitors afford several key advantages for tabletop research and deployment which makes their use a compelling option over projector-based systems. Generally, LCD monitors provide a higher display resolution than projectors for a lower price. For instance, a screen with full 1080p HD resolution will cost several thousands of dollars less than a projector with a similar pixel output. Additionally, the slim profile of LCD monitors makes them easy to embed into the structure of a tabletop, and unlike projectors they don't have issues with key-stoning and throw-distance. There are however challenges that need to be overcome for the multi-touch developer wishing to utilise an LCD screen as a display technology. Firstly, it is imperative to have knowledge of how LCD technology works and be familiar with their manufacturing and assembly.

LCD Technology and Manufacture. This paragraph will briefly describe how LCD's are assembled with a focus on issues that impact tabletop development. The first issue that must be understood is that each pixel of an LCD monitor is comprised of three electronically controlled filters (red, green, and blue filters) which modulate over a backlight to emit a certain colour. Essentially, the LCD glass panel is transparent when no current is running through the screen. Next, on the front and back side of the glass panel are criss-crossing polarizing filters. The polarizing filters give an LCD its black appearance since their opposing orientation blocks visible light. However, polarizing filters do not polarize light within the IR spectrum. So while a LCD panel looks opaque to our eyes, IR light can be transmitted through the screen unperturbed. This concept is crucial for the use of LCD screens in optical multi-touch systems. The next part of the LCD assembly is the back-light and filter-chain. A back-light is necessary in order to illuminate the LCD pixels. The backlight for monitors that are less than 23" consist of a long thin fluorescent light bulb which lines the length of the monitor. Attached to the bulb is an acrylic sheet (called the light guide) which has a honey-comb pattern of white dots. Based on the principle of total-internal reflection, the light from the fluorescent tube travels inside the acrylic sheet until it reflects off the white dot. This method for back-lighting allows for thin displays. For LCD monitors that are 27" and larger, the acrylic-guide method is

replaced by a rail of lights. Behind each screen are rails of thin light fluorescent light bulbs which provide the back-lighting. However, because of the polarizing filters and the method in which the crystals distort and filter light, having only a back-light is somewhat ineffective for illuminating the display. This could be understood if one imagines adjusting their laptop screen in order to achieve the best viewing angle which is orthogonal to their line of sight. Tilting the laptop screen too much and the display image lose much of its color and appearance. To improve the lighting conditions of the display, LCD manufacturers include a layering of several different filters which modulate and affect the backlight in various ways. The most common filters include:

1. Diffuser: this filter diffuses the backlight to disperse in every direction.
2. Fresnel Lens: based on the Fresnel principle, this filter can magnify light with a shorter focal length in different directions. This filter is used to disperse light in 180 degrees.
3. White Reflector: this is an opaque white filter which reflects any light that may have escaped the filters.

These three are the basic filters for LCD monitors. Manufacturers may use more than these filters to improve their product quality; for instance, using different types of diffusers, or more than one Fresnel lens to improve the viewing angle. However, of these filters, the only one that impedes IR light, and is therefore of concern when developing optical multi-touch surfaces is the last white opaque filter. This filter is totally white and hence needs to be removed. All the rest can and should remain to keep optimal viewing performance.

Optical-Based Approaches: Bezel-IR for LCD. There are two broad methods which so far have been successful for creating interactive LCD surfaces with optical sensing. The first, and easiest, is the side-illuminated method of installing IR LEDs around the bezel of the LCD. The LEDs will shine IR light across the top surface of the screen. When a finger touches the LCD screen, light reflects off the finger and traverses through the monitor which is then captured by an IR sensitive camera. The illumination hardware required for this approach is very similar to the FTIR method. Therefore it is often possible to install an FTIR panel on top of an LCD screen and then remove the acrylic; keeping the LEDs intact. With this method, it is recommended to identify IR LEDs with a small package (3mm or SMD) and with a small viewing angle (typical angles for LEDs are 30 degrees, but shorter angles are available at around 15-18 degrees). Choosing a smaller angle will focus the more of the light to shine across the screen. Finally, emerging IR Laser LEDs promise to be the ideal choice for this method because these light sources ensure that the IR light beam is small and focused over the surface [27].

- Advantages**
1. this method is a low-cost approach to multi-touch on a LCD. In fact, it might be the

most cost-effective approach since LCD monitors are cheaper than projectors.

2. no issues with compliant surfaces unlike FTIR method
3. scalable: this method is scalable for larger monitors, such as 32" screens

- Disadvantages**
1. Tangible tracking is difficult to implement because the side-illuminating method cannot illuminate fiducials underneath objects
 2. form Factor: the form factor of this method is dictated by the viewing angle of the camera.
 3. requires more filtering (high-pass) to eliminate false touches
 4. sensitive and gentle dismantling of an LCD is required.

Optical-Based Approaches: Matrix of IR Transceivers.
The second method for enabling multi-touch with LCD screens is to create a matrix of IR transceivers behind the LCD panel as described in [18]. Each transceiver consists of an IR emitter, and an IR detector. The emitter will pulse IR light at a certain frequency which the sensor can detect (similar in theory to IR remote controls except here the light is not encoded to pulse information). When a finger or an object touches the screen, the finger reflects back the light which is detected by the sensor. By creating a matrix that consists of many of these transceivers, it is possible to cover the entire surface area of the LCD screen. The amount of transceivers, their size and pitch (distance between sensors) determines the accuracy and resolution of the touch surface.

- Advantages**
1. this method allows for thin form factors which are comparable to standard LCD screens since the sensors do not require a throw-distance.
 2. this method affords tangible tracking because the sensors can distinguish a fiducial pattern.

- Disadvantages**
1. not easy to self-construct. This method requires expert knowledge of electronics, circuit design, and digital-signal processing (DSP). Also, in order to achieve a high-resolution, the sensor boards should be surface-mounted which gives another barrier to construction since it cannot be handmade.
 2. not scalable: Larger surfaces require more sensors, which increases cost and latency. Covering a 32" screen for example, may require thousands of sensors and a fast-processor to read through all the sensors.

3. BYO MULTI-TOUCH SURFACE

In this section we describe the hardware that is need to built up a optical multi-touch system as described in . We also briefly present different tracking libraries that have to date been used by research groups developing multi-touch hardware and applications.

3.1 Hardware

The hardware required to build an optical multi-touch system includes: infrared illumination sources, silicone compliant surfaces, projection screens, cameras, filters, and projectors.

3.1.1 Infrared Illumination

A common trait of the optical sensing surfaces described previously is that they require an infrared light source. Achieving the right infrared illumination means understanding the different methods of illuminating a surface (DI and DSI) and different the types of IR LEDs (5mm, 3mm, SMD) that are available on the market. We briefly describe how to select and build a suitable IR illuminator for a table design and how to use more advanced circuits using 555 timers or PIC/AVR microcontrollers to pulse the IR light .

Virtually all current IR-based set-ups employ light-emitting diodes (LEDs) as light sources. Being solid-state devices, LEDs have several advantages, such as low heat dissipation, high switching frequency and narrow emission wavelengths. One drawback is that they are significantly more expensive than traditional light sources such as halogen bulbs. In situations where a high continuous light output is desired at a low price, a halogen bulb with a suitable band-pass filter can be an adequate alternative.

Two types of IR LEDs which are commonly used are Osram SFH4250 (SMD) and Osram SFH485 (5 mm). Whether SMD devices or standard LEDs are more appropriate depends on a number of factors. The LEDs have to be mounted to the rim of an acrylic glass plate, which is easier for SMD, as it is possible to simply glue them to the rim with instant glue. After hardening, instant glue is chemically identical to acrylic glass and is therefore able to create a very strong, transparent bond. Mounting standard LEDs requires holes to be drilled into the material, which can be a time-consuming and error-prone process; however wiring and soldering of standard LEDs is far easier. Ideally, LEDs should be powered by a constant-current source; however, in most cases, a constant-voltage source is sufficient. Several LEDs are usually connected in series and powered by a common voltage source. However, the output voltage has to be divided by the number of LEDs in one series; the resulting voltage at a single LED and should match the suggested continuous operating voltage from the datasheet.

Pulsed Illumination. One big problem for the systems described so far is their sensitivity to ambient light from the environment. This problem can be mitigated by adding a small electronic circuit to the set-up which supplies short high-current pulses instead of a continuous low current. The pulse current is usually set high enough such that under sustained operation, the LEDs would be likely to suffer permanent damage after a few seconds. Typically, these pulses are given a duration of between a hundred microseconds and a few milliseconds. The high current levels, which is possible during the short pulse, results in a much higher light output.

The pulse duration and the following cool down period should be kept as close to the specification as possible to prevent overheating of the LEDs. As modern computers are usually

not equipped with the hardware or software to undertake such real-time control tasks, we suggest using a simple microcontroller (e.g., PIC or AVR) or the venerable 555 timer for pulse generation the pulses. A second-level switching element is also necessary, to handle the the high currents which flow through the LEDs. Field-effect transistors (FETs), such as the IRF512 logic-level FET, are particularly easy to integrate with logic circuits and we suggest using these as second-level switches. A final precaution against LED damage is an ordinary fuse. A fuse with a *lower* rating than the expected pulse current should be inserted in series with the LEDs. Should the pulse generation fail for any reason, thereby continuously powering the LEDs, the fuse will blow immediately, thus protecting the other parts of the circuit. Although more current will flow through the fuse than it is rated for, it is unlikely to blow during pulsed operation, as the single pulses are too short. In Figure 10, we present an example circuit diagram which uses the concepts mentioned above. This consists of a total of 5 components, excluding LEDs and power supply, and can easily be built on a bread-board. The two supply voltages of 5 and 12 Volts can be drawn from a standard PC power supply. For details on how to calculate the capacitor/resistor values, see [39].

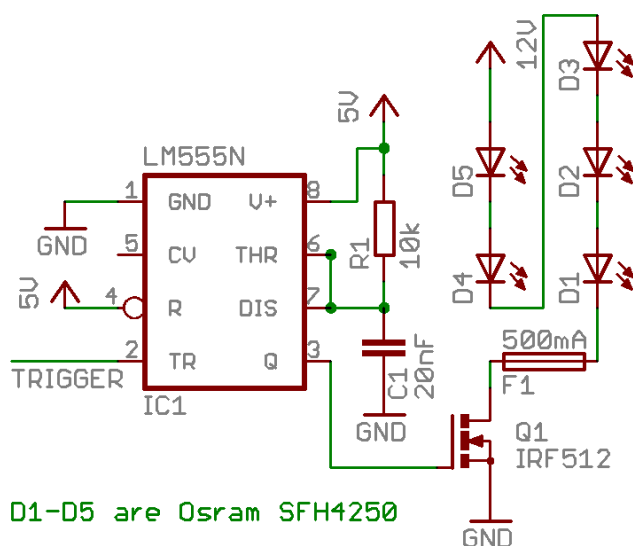


Figure 10: Pulse Generation Circuit

Camera Synchronization. Although pulsing the LEDs already increases total light output, this in itself is not sufficient to gain the significant contrast boost desired with respect to ambient light from the environment. In addition, the pulses need to be synchronized with the camera in such a way that: (1) one pulse is emitted for each camera frame, and (2) each pulse's duration is equivalent to the camera's exposure time.

As the LEDs are usually brighter by approximately one order of magnitude during the pulse, the contrast ratio with respect to environment light is also significantly higher. If the camera exposure time is longer than a single pulse, stray light from the environment is accumulated during the cool

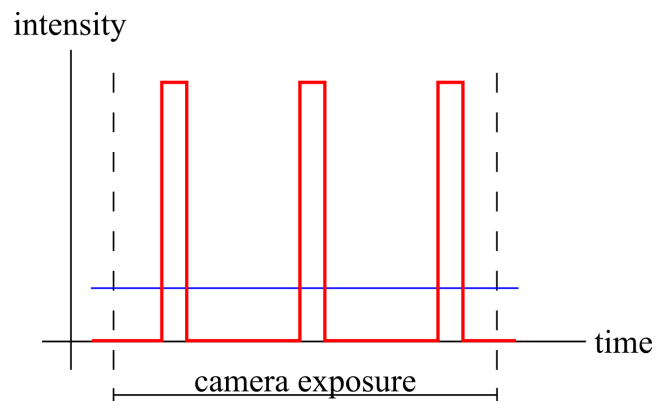
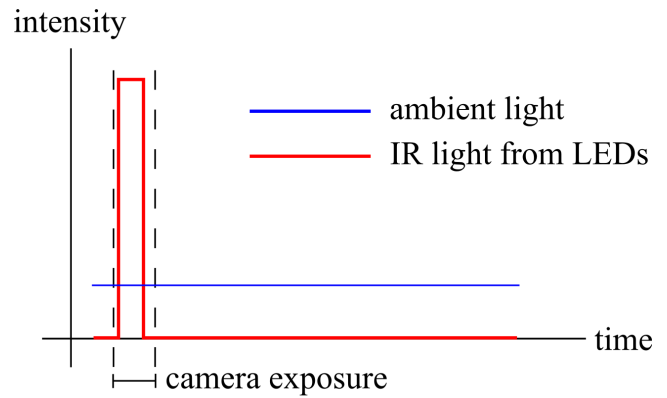


Figure 11: Light intensity per time

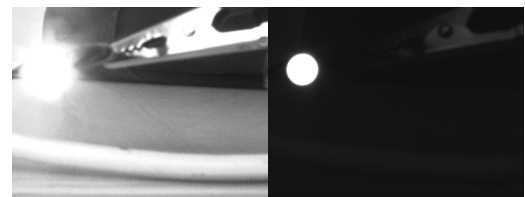


Figure 12: Images from continuous (left) and pulsed (right) operation.

down period between pulses, thereby decreasing the contrast ratio again. In Figure 11, two different operating modes are shown in terms of light output per unit of time. The lower graph shows a mode with several pulses per camera frame; however, in the cool down period a significant amount of environment light is integrated by the camera. In the upper graph, the short exposure time allows only a single LED pulse to be accumulated along with a small amount of stray light.

Figure 12 shows a comparison of the resulting images for the continuous and pulsed operation modes. A single LED is viewed head-on by the camera. In both images, the LED is displayed with maximum brightness (255 in 8-bit mode). However, in the continuous mode, the brightness of the background is approximately 160, whereas in the pulsed mode, the background values are approximately 20, an eight-fold difference.

To realise pulsed operation mode, the camera needs a configurable trigger output and exposure duration. However, these are standard features incorporated in almost all industrial-grade cameras. Some camera models even allow the generation of the entire control pulse with the trigger output, thereby reducing the external circuitry to 2 components (FET and fuse).

For illustrative purposes we can consider how to calculate the correct pulse/exposure duration for a specific camera and LED combination (Pointgrey Firefly MV and Osram SFH4250 LEDs) – for more details of the component characteristics see [31]. If we assume a frame rate of $f = 60\text{Hz}$ then one full pulse/cool down cycle must have a duration of $D_{max} = \frac{1}{f} = 16.67\text{ms}$. If we are operating the LEDs at a voltage of 2.4 V (12 V divided by 5 LEDs) then the current is 1 A. We now have to calculate the total cycle duration, based on the duty cycle for each curve and the allowed pulse duration at a current of 1 A. For example, at a duty cycle of 3.3%, the pulse duration is approximately $t_P = 120\mu\text{s}$ for a total cycle duration of $D = 3.6\text{ms}$. At a ratio of 1% with a pulse duration of $t_P = 250\mu\text{s}$, the total duration already rises to $D = 25\text{ms} > D_{max}$, which is too long. We must therefore select a duty cycle of 2%, resulting in a pulse duration of $t_P = 200\mu\text{s}$ with a total duration of $D = 10\text{ms}$, which still offers a comfortable safety margin. Of course, the camera must be able to provide such short exposure times (as is the case for the Pointgrey Firefly MV).

3.1.2 Cameras, Lenses, Filters, and Projectors

Cameras. FTIR and DI rely on cameras to detect fingers touching the surface. The near-IR spectrum is used to discriminate the background and projected imagery from illumination due to the finger tips. Visible light is cut off using special filters. Both camera and filter choice effect the signal quality.

Camera sensors that are capable of detecting IR light are required in FTIR or DI systems but the sensitivity of CMOS/CCD image sensors to infrared light varies considerably. Infrared light often appears as a bright white spot with a blue/dark purple glow. When choosing a camera it is important to find out which sensor is used and whether the data sheets are available for this sensor. Most data sheets contain a chapter with the spectral sensitivity characteristics. Often a graph shows how sensitive the sensor is to specific wavelengths. In many cases illuminators are used which have a wavelength of 880 nm.

For low cost initial prototypes a USB web camera such as the Philips SPC900NC which uses a Sony CCD image sensor (type: ICX098BQ) is ideal. The spectral sensitivity characteristics are displayed in Figure 13.

Web cameras often contain an infrared filter to block ambient infrared light. This filter layer must be removed and in some cases it is designed to be detachable, although often it is either glued on to the lens or applied as a coating on the camera sensor itself. The Philips camera has the infrared blocking filter glued onto the lens, therefore it is necessary to replace the original lens with a new one.

Spectral Sensitivity Characteristics (excludes lens characteristics and light source characteristics)

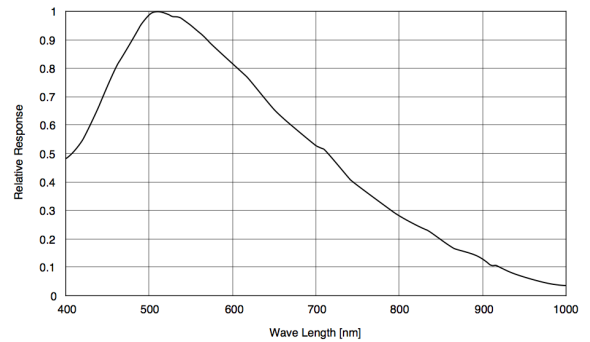


Figure 13: The spectral sensitivity characteristics of the Sony ICX098BL CCD image sensor (excludes lens characteristics and light source characteristics).

Whilst high-end consumer USB cameras are capable of transmitting images of VGA resolution (640×480 pixels) at reasonable frame rates, they often introduce a latency. Any latency will reduce the responsiveness of the multi-touch interface; therefore FireWire based cameras are preferred, e.g. the *Unibrain Fire-i board camera colour* [38]. This camera uses the same sensor (Sony ICX098BQ) as the Philips web camera but has a much lower latency. Depending on the size of the display and the projected image, cameras should normally be run at VGA resolution in order to achieve a reasonable precision. Smooth interaction requires a frame rate that is at least 30 fps.

Because the camera only needs to see infrared illuminated objects, it is advisable to mount an IR band pass filter to prevent distortion from the projected image. For optimal performance this should be a (relatively expensive) band pass filter which matches the IR wavelength of the LEDs although an alternative (cheaper) solution is to use an over-exposed developed negative which acts as a (less specific) IR band pass filter.

To conclude: consumer webcams work well in many cases, but:

- often an IR filter needs to be removed;
- IR sensitivity can be an issue;
- the frame rate is often limited to 15-30 fps.

Alternatively, industrial grade cameras have:

- a higher frame rate (60 to >120 fps);
- special monochrome versions with better IR sensitivity;
- better image quality that reduces the need for image processing;
- ...but they are more expensive.

Lenses, Exposure & Gain and Filters. The exposure time controls how much light reaches the sensor in the camera. Setting the exposure appropriately can be tricky but is important for good tracking results. Gain brightens images and increases contrast, but too much gain can lead to unwanted noise in an image. Integrating wide angle lenses in a system allows smaller distances between the camera and the surface. Lens correction and image rectification waste pixels and so reduce tracking accuracy. Used to cut out background and projected image, cut-off filters work well in controlled lighting conditions. Many materials can be used as DIY filters:

- developed photo film;
- strong sunglasses;
- band pass filters only let a specific wavelength pass (deals better with ambient and scattered IR illumination; creates a higher contrast between contacts and background).

Projectors. When selecting a digital projector, one important factor that must be considered is the display resolution. Depending on the type application a resolution of at least 1024×768 pixels (XGA) is usually required. Additionally, when choosing a specific type of projector—common choices being Digital Light Processing (DLP) or Liquid Crystal Display (LCD)—it is important to consider both the contrast ratio and the number brightness (in lumens) the projector is capable of producing. Since rear projection is widely used in many tabletop interfaces a reduced brightness is often preferable.

Selecting a digital projector suited for a multi-touch display is often more complex than one would imagine. In most cases office projectors are not suitable because of their long throw, which is the distance between the projector and projection surface required to produce a clear focussed image. It is possible to use mirrors to reduce this distance, but this reduces the quality and brightness of the image and significantly complicates the design of the device. It is recommended to use a front surface mirror to remove the double projection (ghosting) that can occur due to the glass front of a conventional mirror.

We have considered several short throw projectors that are currently available on the market. Based on the specifications and prices (see Table 1), the 3M DMS 700 seemed the best choice. The 3M DMS 700 is capable of projecting a screen size with a diagonal of 102 cm from a distance of 50 cm.

3.1.3 Silicone and Projection surfaces

In this section we share our experiences of different layers required to create a functional FTIR surface. These include compliant surfaces, various projection screens, and film to block interference from IR light above the table surface.

Compliant Layer. The simplest way to get started with FTIR is just to use a layer of polycarbonate augmented with

Projector type	Native resolution	MSRP
3M DMS 700	1024x768	\$ 2765
JVC DLA-SX21SU	1400x1050	\$ 11995
NEC WT610	1024x768	\$ 3999
Sanyo PLC-XL50	1024x768	\$ 2900
Toshiba TDP-ET20U	854x768	\$ 999
Toshiba TDP-EW25	1280x800	\$ 1899

Table 1: Overview short-throw digital projectors. Including the manufacturers suggested retail price (MSRP) July 2008.

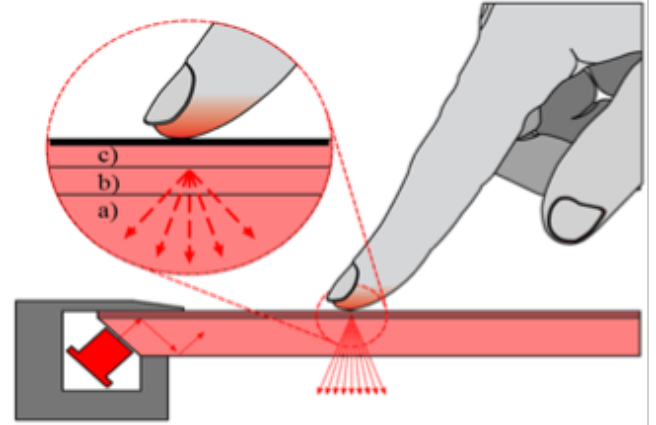


Figure 14: The three layers needed to track the finger touches: the polycarbonate plate (a) is covered with a compliant surface layer (b) and a diffuse layer (c) on top.

a frame of infra-red LEDs. However, with this configuration users often must press hard on the surface in order to trigger the FTIR effect. Additionally, when dragging a finger on the surface, such as when performing a motion gesture, friction may decrease the intensity of the FTIR effect caused. Therefore, many researchers use an additional layer (a compliant surface layer) on top of the polycarbonate material to improve the sensitivity of the surface. These compliant surfaces are typically composed of a soft and transparent material which is placed between the polycarbonate sheet and the diffuse (projection screen) layer. Figure 14 highlights the relevant layers of a commonly used composition. When pressure is applied on the surface, the coupling of the diffuse layer and the polycarbonate surface triggers the FTIR effect; this effect is intensified by the compliant surface layer.

Finding the correct material for a compliant surface is crucial. When experimenting with different materials we noticed two different problems that can occur with the layer: either it does not set off a strong-enough FTIR effect (see Figure 15 (a)), or it sticks to the surface, constantly triggering the FTIR effect even after a finger has been removed (see Figure 15 (b)). In the worst case, if the wrong combination of compliant material and projection layer is used; the two layers can stick together permanently.

The best results for the compliant surface were achieved with SORTA-ClearTM40 [36] and ELASTOSIL®RT 60 [40] silicone, both materials being relatively hard (Hardness Shore

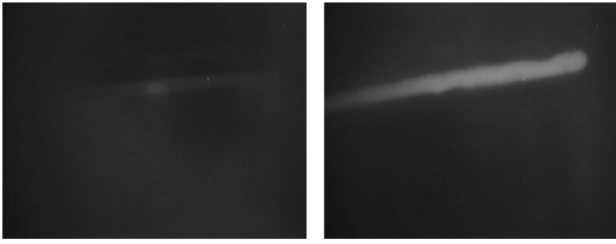


Figure 15: Using the wrong combination of materials can result in two main problems: (left) either the FTIR effect is not strong enough; or (right) the layers stick together.

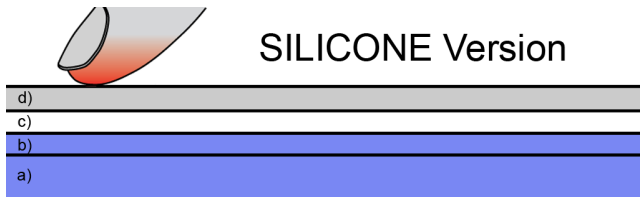


Figure 16: Silicone compliant layer: the gap (c) is between the projection surface (d) and the combined silicone (b) polycarbonate (a) layer.

A ≥ 40), non tacky and very clear. Once hardened, both silicone layers can easily be removed from, and re-attached to, the polycarbonate surface. However, using silicone as a compliant surface poses one problem as the material comes as a gel which must be poured evenly over the surface. ELASTOSIL®RT 601 is less viscous and hence easier to pour, resulting in fewer bubbles in the vulcanized layer.

As an alternative to silicone, we found that a thin layer of latex also works well. This also has the significant advantage of not having to be poured, reducing the construction time for the combined layer significantly. Furthermore, latex is easier to handle, cheaper to produce, and more readily accessible as an off-the-shelf component. Moreover, latex does not stick to neighbouring layers, as with other alternative compliant surface materials, so latex can be combined with a wider variety of projection layers. In contrast to silicone, the latex must be combined with the projection layer; with an air gap between the latex and the polycarbonate base plate. In the silicone version we have exactly the opposite requirements. Figures 16 and 17 show this difference between the latex and silicone layer construction.

Projection Layer. Depending on the compliant material (silicone or latex) it is possible to use different materials as a projection screen. The main requirements are that an air gap should be achievable between two layers and that screen allows the triggering of the FTIR effect. Not all materials meet these requirements. Figure shows different results for projection materials on top of silicone.

Figure 18 (a) shows an optimal result for FTIR with a high contrast touch point. Materials that resulted in too dark

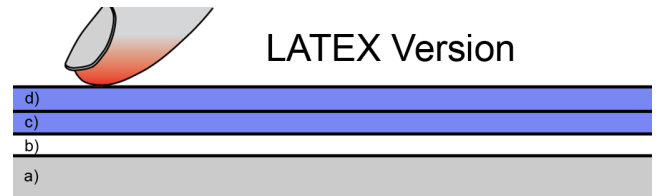


Figure 17: Latex compliant layer: the projection (d) and the latex layer (c) must be combined; the gap (b) is between these two and the polycarbonate plate.

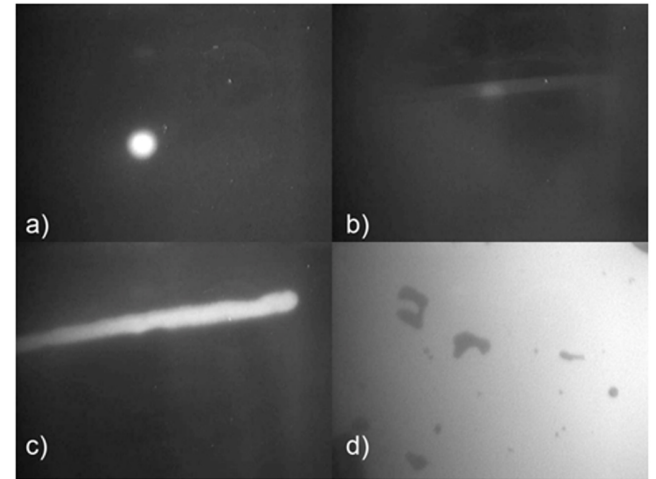


Figure 18: (a) Rigid PVC (backlit) (b) Rosco translucent [4] (c) Sihl polyester film 100 mat, and (d) HP backlit UV.

touch points (b) or showed permanent traces on the silicone as well as materials that completely stuck to the silicone (c) are not suitable for FTIR. Rigid PVC and tracing paper appear to be a good solution in combination with silicone. They do not stick to the silicone but trigger the FTIR effect quite well.

For the latex version, we found HP Colorluculent Backlit UV to be a good choice. HP Colorluculent Backlit UV foil was originally designed for use in backlit signs. Similar to rear-projection screens it yields a diffuse image without any hotspots from the projector, making it a good rear-projection surface. Because of its glossy backside, it cannot be used with the silicone, because it sticks to the silicone as shown in Figure 18 (d). Rosco screens can also be combined with latex and the latex sticks well to the screen.

3.1.4 Hardware Integration

On this section we present different methods for combining all the “ingredients” together to get your multi-touch surface ready to use. The most trivial set-up is an interactive wall, where the camera, and projector are placed orthogonally behind the interactive surface (e.g. [35]). This setup requires a lot of space behind the multi-touch surface depending on what kind of hardware (projector, camera) is used (see figure 19). A multi-touch console such as this can be built without any mirrors as shown in figure 19 b).

FTIR Pipeline

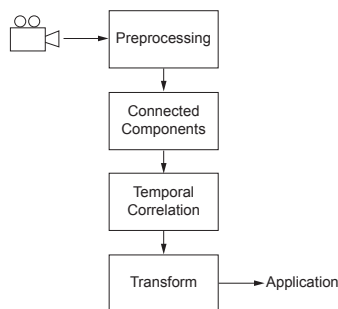


Figure 21: FTIR Tracking Pipeline

For constructing and designing a interactive tables mirrors commonly have to be used to allow a camera and projector configuration which results in a usable table height. Figure 20 shows this setup for both FTIR or DI based surfaces. A example construction sheet can be found online¹. Of course multi-touch surfaces are not limited to being rectangular. Some example of other shapes surfaces can be found on <http://hci.rwth-aachen.de/pingutouch>, where a hexagonal table is presented, or in [2], where a round table is described.

3.2 Software

Tracking touches on a surface involves setting up a pipeline of image processing operators that transform a camera image into user interface events.

3.2.1 FTIR Tracking Pipeline

Figure 21 shows the typical imaging pipeline of an FTIR set-up. Images captured by a camera are first pre-processed to remove any unchanging parts using history subtraction. A connected components algorithm (described e.g. in [15]) finds bright regions in the pre-processed image. These are the areas where something is touching the surface. Post-processing involves finding corresponding touches in different camera frames (temporal correlation) and transforming the camera coordinates to screen coordinates.

3.2.2 DI Tracking Pipeline

DI tracking is a more complex process but allows for proximity as well as touch to be sensed. DI Touch detection exploits the fact that objects at a distance from the surface are appear blurred. *reactTable* [21] does this by adaptive thresholding based on the curvature of the luminance surface (see [6] for a detailed description of the algorithm). The multimedia platform *libavg* [51] used in the *c-base MTC* pioneered the use of a high-pass filter to achieve the same effect.

Figure 23 shows images generated in a typical DI tracking pipeline. As can be seen, it splits the image pipeline and

¹https://www.libavg.de/wiki/images/4/46/Mtc_construction.pdf

DI Pipeline

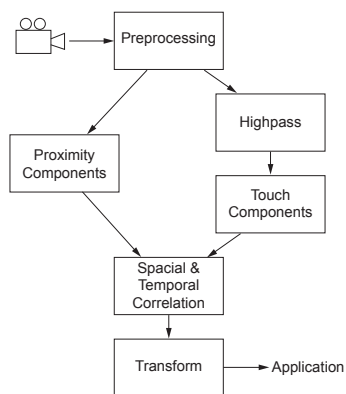


Figure 22: DI Tracking Pipeline

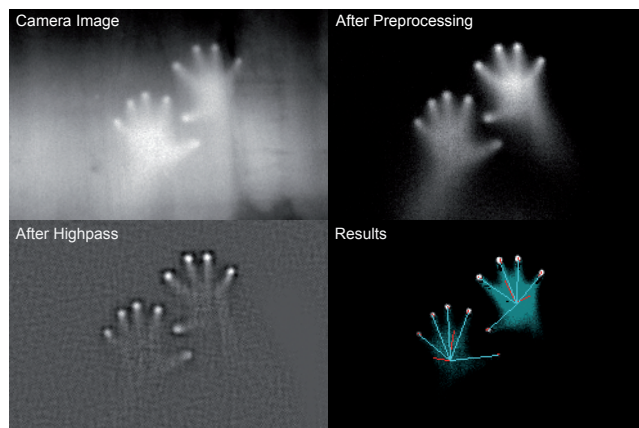


Figure 23: Sample Images of the DI Tracking Pipeline

runs the connected components algorithm twice, once each for touch and once for proximity sensing. Touch sensing involves an additional high-pass filter to isolate areas very close to the surface. After the regions have been found, touch and proximity information can be correlated. The bottom right image in Figure 23 shows the result of this process: Fingers touching the surface have been identified and associated with hands.

Figure 23 illustrates the output of *libavg* image processing.

3.2.3 Interface Considerations

The tracking pipeline provides higher level software layers with information about finger and hand positions. TUIO [22] uses Open Sound Control over UDP to transmit this information, but is rather low level. By default Touchlib and many other libraries come with a wrapper which sends TUIO events over the commonly used OpenSound Control² protocol. For many modern programming languages such as C#, Adobe Flash (Actionscript 3), Java, Max/DSP, Processing, Pure Data, Python and Visual Basic, OSC libraries

²OSC
OpenSoundControl/

<http://www.cnmat.berkeley.edu/>



Figure 19: a) Multi-touch wall setup, multi-touch console without mirrors, sketch of a multi-touch table with mirror (projector is mounted in the small tower that can be seen in the image).

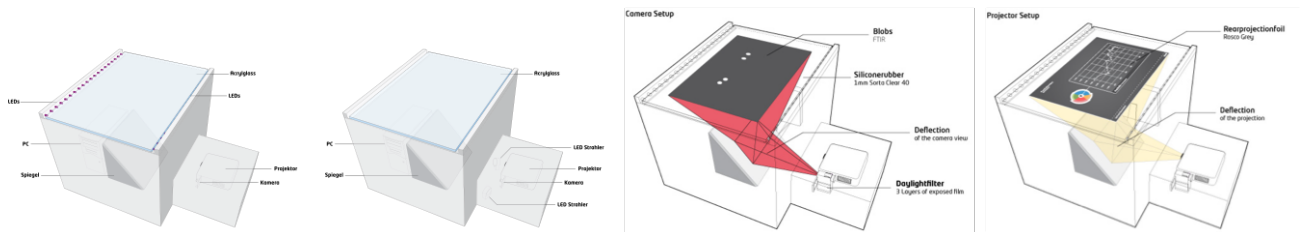


Figure 20: General multi-touch table setups for DI and FTIR with projector and camera.

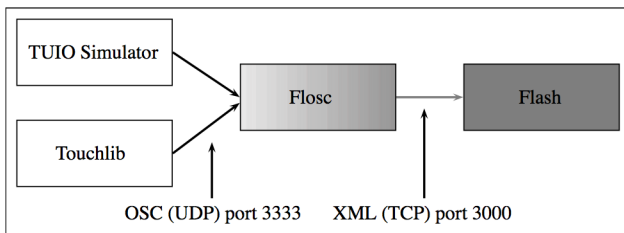


Figure 24: Schematic view of sending TUIO events over the OSC protocol. Flosc converts the UDP packages to TCP packages.

are available. When using Flash it is required to convert UDP packages to TCP. This can be done by using the tool Flosc which acts as a proxy (Figure 24).

Work is in progress to provide higher-level interfaces (*libavg* [51], *libtisch* [10]). *libavg* which includes event processing that correlates touches to a hierarchy of on-screen widgets³. This corresponds to the mouse event handling that window systems provide and hence affords the basis for robust implementation of classical GUI widgets like buttons and scrollbars. Both libraries support emerging gesture standards that allow for dragging, rotating and scaling of GUI elements.

³https://www.libavg.de/wiki/index.php/Event_Handling

When an application uses the OSC protocol, it is only be able to receive events containing properties of the detected blobs. It is not possible to adjust the settings of Touchlib from the application. However, since OSC uses the UDP network protocol to transfer data it makes it possible to create a set-up in which a dedicated system provides blob tracking and transfers the data to another system which provides the visualization.

At higher levels, window-system-like event processing, classical GUI widgets (buttons etc.) and emerging gesture standards (dragging, rotating and scaling elements, for instance) are supported by some libraries.

libavg *libavg* [51] is a multimedia platform that includes support for object tracking using cameras, including DI and FTIR tracking. Tracking is partially implemented on the GPU for performance and robustness. Since the tracking is tightly integrated with a hierarchical layout system, *libavg* is able to provide a very high-level interface to touch events (See section 3.2.3 for details). It is also the only library in this list that supports the full DI pipeline including hand-finger correlation. For compatibility, a TUIO wrapper is also provided. *libavg* supports Linux and Mac OS X. A windows port is in progress.

Multi-touch lib T-Labs Another library is the Java multi-touch library⁴ developed at the *Deutsche Telekom Lab-*

⁴<http://code.google.com/p/multitouch>

oratories. This library is released under the GNU Public License. It contains a set of common algorithms designed to work with any multi-touch system such as routines to label connected components and track features. By using an application layer, it is easy to manipulate objects and transform (position, rotate, scale) them. The library also comes with a module for accessing cameras such as the PointGrey Dragonfly2 and the library is able to stream the data into the aforementioned TUIO protocol. The main idea of the library is to provide common set algorithms, not wired to a special category of multi-touch devices:

- Feature Layer**
- Labeling (Connected Components)
 - Tracking (Features)
 - Physical Moments
- Application Layer**
- Manipulators
 - Transform (Position, Rotation, Scale)

In addition to that the library also provide a minimal module for accessing cameras, Currently Point Grey camera devices are supported utilizing Quicktime for Java and through the PointGrey FlyCapture SDK via the Java Native Interface. (According to the project website Lib-DC-1394 will follow soon). It was used for the project presented in [2].

OpenFTIR OpenFTIR is a framework by David Smith's Group that's under development right now. It's closely tied to Windows APIs, which allows it to perform quite well: A 640x480 pixel frame can be processed in about 1.5 ms. It also computes a motion vector of the touch blobs by simply subtracting the previous frame from the newest frame: The white portion of the blobs show the forward direction, the black portion the backward direction of the movement. Existing example applications provided by OpenFTIR include a photo sharing application, a painting application and a musical application (user can place notes on a bar to create music).

TouchLib [41] is a free open source cross platform multi-touch framework which provides video processing and blob tracking for multi-touch devices based on FTIR and DI. Video processing in Touchlib is done through the Intel's OpenCV graphics library [19]. Touchlib currently runs on MS Windows, Linux and Mac OS X, and the probably most used library by beginners.

VVVV A nice software toolkit for rapid development of prototypes is vvvv. It uses a visual programming paradigm similar to Quartz Composer and Max/MSP and allows for the rapid connection of blob input data with various visualization methods. Unfortunately, it's Windows only (plugins for other visual programming languages are however becoming available, such as a TUIO client for QuartzComposer). vvvv' it is suitable for many tasks involving common computer related media like audio, video and 3d animation. Besides creating and transforming content for the aforementioned media, vvvv is also well trained in receiving input from and generating output to various external devices. The visual programming interface allows users to easily develop their own applications. vvvv also supports Multi-Projection setups, 3D rendering and analysis and physics.

Other libraries Besides the five presented libraries there are dozens of others available on the open source market for example OpenTouch or the Touchkit library.

4. SYSTEM LATENCY

In order to measure the influence of hardware on the performance of a multi-touch device, we measured several parts of a multi-touch system separately. For those parts we did not measure, we give guidelines regarding possible latencies. A complete overview can be found in [29].

4.1 Camera

In the first stage of the pipeline a camera frame is acquired for processing. This includes several steps that introduce latency:

1. The integration time. This is the actual time that the sensor is picking up light. The amount of time this takes is dependent on the exposure setting of the camera.
2. The sensor readout time, that is, the time that is needed to transfer data from the sensor to the camera electronics. For some cameras, this happens parallel to the data transfer to the computer. Others transfer the complete image to an internal memory before beginning the transfer.
3. The transfer time from the camera to the computer over the firewire bus.

We did not measure times for 1) and 2). However, sensor readout and integration cannot happen at the same time, so the maximum latency these two steps can introduce is 33 ms for a camera running at 30 frames per second. The delay caused by the transfer over the firewire bus can be calculated by referencing the Instrumentation & Industrial Digital Camera ⁵ specifications. Each frame is transferred through the FireWire bus. The time it takes to transfer a frame depends on the format at which the camera operates. The FireWire bus sends out packages in 125 microsecond time intervals. The camera operates at 8 bit monochrome with a resolution of 640x480 pixels at 30 frames per second. According to the IIDC specifications this format uses 240 packets per frame (1280 bytes per packet). The time needed for each frame to be transmitted to the system is $240 \times 125 \mu\text{sec} = 30$ milliseconds.

4.2 Touchlib

In the next stage of the pipeline the camera image is processed by Touchlib (see above).

4.2.1 Latency measurement method

In order to measure the latency of Touchlib, timers were added to the Touchlib source code. Each image filter was measured separately. All tests were done on the same machine used for the experiments.

⁵IIDC <http://damien.douxchamps.net/ieee1394/libdc1394/iidc>

4.2.2 Latency results

During the implementation of the timers in the source code a ‘bug’ in Touchlib was discovered which influenced the performance of Touchlib. After each image processing and blob tracking loop, the loop stalls for 32 ms (*Sleep(32)*). This means that the actually processing time of Touchlib is the total Touchlib time from Table 2 plus 32 ms resulting in an average of 65 ms.

4.3 Application

After touchlib, the application needs to act on the data. Besides the actual application processing time, which is of course application-dependent, this causes an operating system context switch if a network protocol such as TUIO is used. With most current operating systems, an additional latency in the order of 10 ms is introduced by the context switch, with a lot of variation possible depending on the system configuration and overall load.

4.4 Digital projector

In the last stage of the pipeline the resulting image displayed using the digital projector. An earlier paper by Woods et al. [49, 48] found a negligible delay of less than one millisecond using DLP projectors. We experienced an easily noticeable delay using one projector, however, and did some measurements of our own.

4.4.1 Latency measurement method

Based on the latency measurement tool described in the paper [34] a re-implementation was created.

The measurement tool allows us to compare the number of frames the digital projector lags compared to the reference CRT monitor. In order to use the tool it is required to turn on vertical synchronization in the display settings. For our measurements we used an Iiyama Vision Master Pro 510 as reference CRT monitor. When starting the tool it is required to input the native resolution and refresh rate of the digital projector. The 3M DMS 700 uses a resolution of 1024×768 at 60 Hz. While the tool is running a large green bar will move position each time a frame is being refreshed. In our case that means 60 positions per second.

To perform the measurement the display of the CRT is cloned to the native resolution and refresh rate of the projector. Next the projector is aimed at a projection surface placed next to the CRT monitor. To compare the latency it is necessary to take a photograph that includes the image of the CRT and digital projector in one frame. In order to take this picture, a digital camera was used with the shutter speed set to the refresh rate of the screen. The latency in milliseconds can be calculated by comparing the position of the green bar of both devices (see Figure 25) up to a precision of one frame.

4.4.2 Latency results

We performed measurements on multiple digital projectors. The resulting images were compared on the computer. The image from the 3M DMS 700 (Figure 25) showed a delay of six frames. Since the refresh rate was fixed at 60 Hz, the refresh time of a single frame is equal to $\frac{1}{60}$ seconds ≈ 16.667

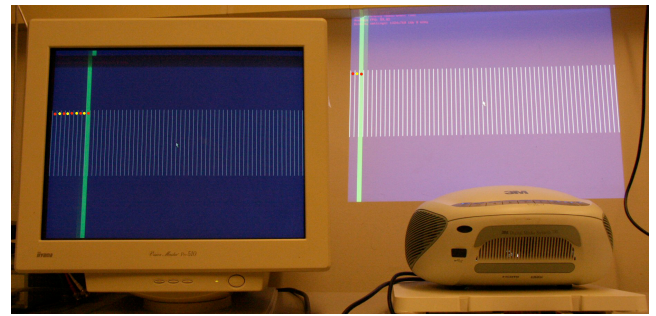


Figure 25: Comparing the latency of the digital projector with a CRT monitor using the latency tool.

ms. By multiplying this value with the number of (delayed) frames we can find the following latency:

$$\frac{1}{60} \times 6 = 0.1 \text{ seconds} = 100 \text{ milliseconds.}$$

After processing the results of the other digital projectors we found out that only the 3M DMS 700 had such a high latency. After contacting 3M support, they informed us this was probably due to the image improvement chip (Hollywood Quality Video, HQV). According to 3M it was not possible to turn off this image processing chip. Results of projectors tested by us are listed in Table 3.

Note that all other projectors measured have a latency within the measurement precision of one frame and that the Epson digital projector has a negative relative latency value. Benchmark results show that the Epson is almost one frame ahead compared to the CRT monitor.

4.5 Total system latency

It is now possible to calculate the total system latency. The results are displayed in Table 4. These test results do not include the application processing time. We assume a total integration and sensor readout time of 20 ms and a context switch time of 10 ms.

Because of the 3M DMS 700 projector and the aforementioned Touchlib bug the total latency is almost 225 milliseconds.

5. CURRENT WORK AND PROJECTS

In this section we highlight ongoing work and useful websites and communities where users can find additional information. This collection of links is unsorted and in no particular order – our selection does not attempt to include all webpages describing self-built multi-touch surfaces.

Microsoft Surface The Microsoft Surface, is the multi-touch product from Microsoft which is developed as a software and hardware combination technology that allows a user, or multiple users, to manipulate digital content by the use of natural motions, hand gestures, or physical objects [46]. More information, videos and gallery available on their webpage <http://www.microsoft.com/surface>. The setup is a slightly modified DI setup

and explained on the (non-technical) website⁶. With Microsoft Touch Wall⁷ and the Microsoft Touch Sphere⁸ Microsoft recently presents two new interactive surfaces. TouchLight <http://www.youtube.com/watch?v=F1br3fDZyUQ> [47]. LucidTouch [43] is also an interesting project done by Microsoft research.

Jeff Han: Multi-Touch Interaction Research As can be seen in figure 2 the work [13, 14] of Jeff Han has a great impact on the community. More information about his multi-touch project (<http://cs.nyu.edu/~jhan/ftirtouch>) and his company can be found online (<http://www.perceptivepixel.com>).

180 180 is an interactive multi-touch surface (DSI) designed by Tim Roth for consulting situations. Traditionally, consulting situations involve two participants sitting on opposite sides to of a table each other. This makes the use of digital media difficult, since it has to be made accessible for both participants. The basis for the project is a typical consulting situation for a private bank from Zurich. The video⁹ shows the application in use. The webpage shows nice graphics of the setup of his table <http://timroth.de/180>. It is a follow up project of tangent¹⁰ also developed at “Hochschule für Gestaltung und Kunst“ Zurich.

c-base MultiTouchConsole MTC The c-base MultiTouchConsole MTC is a DI multi-touch table that was first presented by a group centered around Ulrich von Zadow at the 23rd Chaos Communication Congress in 2006. It is based on the open-source library libavg and served as a test-bed for it’s the multi-touch features. The hardware setup and material needed to build a similar table are well-documented (Web page¹¹, german video¹²), as is the libavg tracking setup¹³. The c-base is a non-profit association in Berlin, Germany that provides an open space for technological experimentation. The MTC has accordingly been used as a basis for interaction experiments such as Jens Wunderling’s LoopArena sequencer [50] and sponc [16], a pong-like game, created by Martin Heisterman.

The Archimedes MTT is a commercial version of the MTC¹⁴. It is sold as a complete self-contained table with high-quality 60 Hz tracking and 3500 Lumen high-contrast projection.

⁶MS Surface: <http://arstechnica.com/news.ars/post/20070530-what-lurks-below-microsofts-surface-a-qa-with-microsoft.html>

⁷MS Touch Wall: <http://www.microsoft.com/presspass/events/ceosummit/default.mspx>

⁸MS Touch Sphere: <http://research.microsoft.com/~benko/projects/sphere>

⁹Video 180: <http://www.timroth.de/180/page3/files/page3-1010-pop.html>

¹⁰<http://zima.ch/tangent/>

¹¹https://www.libavg.de/wiki/index.php/Building_a_MultiTouch_Console

¹²Video mtc: ftp://dewy.fem.tu-ilmenau.de/CCC/23C3/video/23C3-1512-de-homegrown_interactive_tables.m4v

¹³https://www.libavg.de/wiki/index.php/Tracker_Setup

¹⁴www.archimedes-solutions.de/en.html

Pascal Schmidt German “Jugend forscht” Pascal Schmidt, a student at a Germa high school (18), won the technical & scientific youth-contest in Germany 2008 with his home grown DI multi-touch table. The prototype was presented together with a company called mehr:wert at CeBit 2008 in Early 2008. The web pages <http://multitouch.sourceforge.net> and <http://mehrwert.cc/> containing more information on this project.

Fraunhofer Multi-Touch The Fraunhofer multi-touch table was also shown at Cebit 2008 in Hannover (<http://a4www.igd.fraunhofer.de/projects/48/>) demonstrating the possibilities of multi-touch in a planning scenario [20].

TISCH The central element of TISCH is a multi-touch table (TISCH: ¹) that provides room for 4 to 6 concurrent users. A frosted glass table of about 1.10 x 0.7 m is used as a back-projection surface. It is mounted on a robust aluminium frame which contains a projector, an infrared camera and a computer. An acrylic glass sheet placed on top of the projection area has 70 infrared LEDs (Osram SFH4250, pulsed at 2.4 Volts) attached around its rim to provide FTIR-based multi-touch input to the computer via an IR camera.

To gather proximity information, the goal of the creators of TISCH was to create distinct shadows of objects on and above the surface. Therefore they mounted an additional infrared light source at the ceiling above the table. While this increases the complexity of the system and reduces its mobility, interactive tables tend to be stationary equipment which could be integrated into existing conference room tables or placed in a public area as an information booth. For a schematic of the entire hardware setup, see Figure 26.

NUI Group The Nui group – or Natural User Interface group – is one of the most active groups dealing with multi-touch interaction and interactive media. Their goal is it to create open source machine sensing techniques which will benefit artistic and educational applications. Their focus is on “Open Source Interface”, which is solely for accelerating development of existing hardware and sensing solutions. On their webpage <http://nuigroup.com> they provide a wiki and a active forum. TouchLib (see software section), TouchEvent <http://touchevent.riaforge.org> and OpenTouch <http://opentouch.info> are their three main projects.

Guides Other useful guides and wikis can be found on: FTIR Multitouch and Display Device¹⁵, David Smith and David Holmans Guide about “Building a Multi-Touch Sensitive Table”¹⁶, Harry van der Veens website¹⁷ with working Flash applications for multi-touch applications¹⁸ or Johannes Schönings multi-touch pod-

¹⁵Tangible Interaction Surface for Collaboration between Humans

¹⁵<http://www.lowres.ch/ftir>

¹⁶http://dundee.cs.queensu.ca/wiki/index.php/Building_a_Multi-Touch_Sensitive_Table,

¹⁷http://www.multitouch.nl/documents/multitouchdisplay_howto_070523_v02.pdf

¹⁸<http://www.multitouch.nl/?p=23>

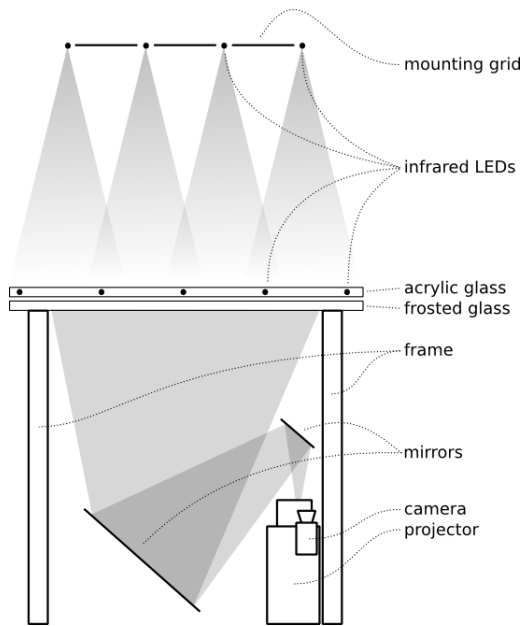


Figure 26: Hardware Setup hardware setup

cast¹⁹.

Other useful links In an unsorted collection:

<http://naturalinteraction.org/>
<http://www.i-bar.ch>
<http://www.mindstorm.com/>
<http://xenakis.3-n.de>
<http://touchkit.nortd.com>
<http://nortd.com/cubit>

6. UPCOMING EVENTS

Mid October 2008, UIST will be held in Monterey, USA. The program includes multi-touch related work which will be presented there in the “Touch and Pressure ” session. In April 2009 a workshop on “Multitouch and Surface Computing Workshop” will be held at CHI 2009 in Boston, USA. And of course we hope to see you all at IEEE Tabletops and Interactive Surfaces in 2009.

7. ACKNOWLEDGMENTS

We would like to thanks Peter Konopatzky, Sven Reißig, Michael Scholz, Antonio Krüger and David Smith for their comments, input, help and support for our Bootcamp at the IEEE Tabletops Workshop 2008 in Amsterdam.

8. REFERENCES

- [1] Apple Inc. Apple iPhone. <http://www.apple.com/iphone>. [Online; accessed 23-September-2008].
- [2] K. Bredies, N. A. Mann, J. Ahrens, M. Geier, S. Spors, and M. Nischt. The multi-touch soundscape renderer. In *AVI '08: Proceedings of the working conference on Advanced visual interfaces*, pages 466–469, New York, NY, USA, 2008. ACM.
- [3] B. Buxton. Multi-Touch Systems that I Have Known and Loved. <http://www.billbuxton.com/multitouchOverview.html>. [Online; accessed 23-September-2008].
- [4] W. Buxton and B. Myers. A study in two-handed input. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 321–326, 1986.
- [5] E. Costanza and J. Robinson. A region adjacency tree approach to the detection and design of fiducials. *Vision, Video and Graphics (VVG)*, pages 63–70, 2003.
- [6] E. Costanza, S. Shelley, and J. Robinson. Introducing audio d-touch: A tangible user interface for music composition and performance. *Proc. of the 6th Int’l Conf. on Digital Audio Effects (DAFX)*, pages 63–70, 2003.
- [7] P. Dietz and D. Leigh. DiamondTouch: a multi-user touch technology. *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 219–226, 2001.
- [8] R. Downs. Using resistive touch screens for human/machine interface. Technical report, Texas Instruments Incorporated, Texas Instruments, Post Office Box 655303, Dallas, Texas 75265, 2005.
- [9] R. G. J. et al. US patent #3,673,327: Touch actuable data input panel assembly, 1972.
- [10] Florian Echtler. TISCH: Tangible Interactive Surfaces for Collaboration between Humans. <http://tisch.sourceforge.net/>. [Online; accessed 23-September-2008].
- [11] Google Inc. Google Insights for search. <http://www.google.com/insights/search>. [Online; accessed 23-September-2008].
- [12] Google Inc. Google Trends. <http://www.google.com/trends>. [Online; accessed 23-September-2008].
- [13] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 115–118, New York, NY, USA, 2005. ACM Press.
- [14] J. Y. Han. Multi-touch interaction wall. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Emerging technologies*, page 25, New York, NY, USA, 2006. ACM Press.
- [15] Y. Han and R. Wagner. An efficient and fast parallel-connected component algorithm. *Journal of the ACM (JACM)*, 37(3):626–642, 1990.
- [16] M. Heisterman. sponc. <http://sponc.de>. [Online; accessed 23-September-2008].
- [17] K. Hinckley, R. Pausch, D. Proffitt, and N. Kassell. Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 5(3):260–302, 1998.
- [18] S. Hodges, S. Izadi, A. Butler, A. Rrustemi, and B. Buxton. Thinsight: versatile multi-touch sensing for thin form-factor displays. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 259–268, New York, NY, USA, 2007. ACM.
- [19] Intel. Open source computer vision library: Reference

¹⁹<http://phobos.apple.com/WebObjects/MZStore.woa/wa/viewPodcast?id=268564447>

- manual. 2001.
- [20] Y. Jung, J. Keil, J. Behr, S. Webel, M. Zöllner, T. Engelke, H. Wuest, and M. Becker. Adapting x3d for multi-touch environments. In *Web3D '08: Proceedings of the 13th international symposium on 3D web technology*, pages 27–30, New York, NY, USA, 2008. ACM.
- [21] M. Kaltenbrunner and R. Bencina. reactIVision: a computer-vision framework for table-based tangible interaction. *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 69–74, 2007.
- [22] M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza. TUIO: A protocol for table-top tangible user interfaces. *Proc. of the The 6th Int’l Workshop on Gesture in Human-Computer Interaction and Simulation*, 2005.
- [23] L. R. Kasday. US patent #4,484,179: Touch position sensitive surface, 1984.
- [24] J. B. Mallos. US patent #4,346,376: Touch position sensitive surface, 1982.
- [25] N. Matsushita and J. Rekimoto. HoloWall: designing a finger, hand, body, and object sensitive wall. *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 209–210, 1997.
- [26] Microsoft Corporation. Microsoft Surface. <http://www.microsoft.com/surface>. [Online; accessed 23-September-2008].
- [27] N. Motamedi. Hd touch: multi-touch and object sensing on a high definition lcd tv. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3069–3074, New York, NY, USA, 2008. ACM.
- [28] R. E. Mueller. US patent #3,846,826: Direct television drawing and image manipulation system, 1974.
- [29] L. Y. L. Muller. Multi-touch displays: design, applications and performance evaluation. Master’s thesis, University of Amsterdam, Amsterdam, the Netherlands, June 2008.
- [30] Nintendo Co., Ltd. Nintendo DS. http://www.nintendo.co.uk/NOE/en_GB/nintendo_ds_1023.html. [Online; accessed 23-September-2008].
- [31] OSRAM Opto Semiconductors. SFH 4250 Datasheet. <http://www.osram-os.com>.
- [32] J. Rekimoto. Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 113–120, New York, NY, USA, 2002. ACM.
- [33] T. Roth. Dsi - diffused surface illumination. <http://iad.projects.zhdk.ch/multitouch/>. [Online; accessed 23-September-2008].
- [34] S. Rusdorf, G. Brunnett, M. Lorenz, and T. Winkler. Real-time interaction with a humanoid avatar in an immersive table tennis simulation. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):15–25, 2007.
- [35] J. Schöning, B. Hecht, M. Raubal, A. Krüger, M. Marsh, and M. Rohs. Improving interaction with virtual globes through spatial thinking: helping users ask "why?". In *IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces*, pages 129–138, New York, NY, USA, 2008. ACM.
- [36] Smooth-On. Sorta-clear 40. http://tb.smodev.com/tb/uploads/SORTA_CLEAR_40_32707.pdf, note = "[Online; accessed 23-September-2008]".
- [37] Tyco Electronics. Elo TouchSystems. <http://www.elotouch.com/>. [Online; accessed 23-September-2008].
- [38] Unibrain. Fire-i digital board camera. 2008.
- [39] T. van Roon. 555 timer tutorial. <http://www.uoguelph.ca/~antoon/gadgets/555/555.html>.
- [40] Wacker Silicones. Elastosil rt. http://www.wacker.com/internet/webcache/en_US/PTM/TM/Elastosil/Elastosil_RT_Addition/ELASTOSIL_RT_601.pdf.
- [41] D. Wallin. Touchlib: an opensource multi-touch framework. 2006.
- [42] R. M. White. US patent #4,484,179: Tactile sensor employing a light conducting element and a resiliently deformable sheet, 1987.
- [43] D. Wigdor, C. Forlines, P. Baudisch, J. Barnwell, and C. Shen. Lucid touch: a see-through mobile device. In *IUIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 269–278, New York, NY, USA, 2007. ACM.
- [44] Wikipedia. Google trends — wikipedia, the free encyclopedia, 2008. [Online; accessed 25-September-2008].
- [45] Wikipedia. Indium tin oxide — wikipedia, the free encyclopedia, 2008. [Online; accessed 28-July-2008].
- [46] Wikipedia. Microsoft surface — wikipedia, the free encyclopedia, 2008. [Online; accessed 24-August-2008].
- [47] A. D. Wilson. Touchlight: an imaging touch screen and display for gesture-based interaction. In *ICMI '04: Proceedings of the 6th international conference on Multimodal interfaces*, pages 69–76, New York, NY, USA, 2004. ACM.
- [48] A. Woods and T. Rourke. The compatibility of consumer DLP projectors with time-sequential stereoscopic 3D visualisation. *at Stereoscopic Displays and Applications XVIII, San Jose, California, January*, 2007.
- [49] A. Woods, T. Rourke, and K. Yuen. The Compatibility of Consumer Displays with Time-Sequential Stereoscopic 3D Visualisation. In *Proceedings of the K-IDS Three-Dimensional Display Workshop*, volume 2006, page 21, 2006.
- [50] J. Wunderling. LoopArena sequencer. <http://www.sport4minus.de/loopArena/mtc.html>. [Online; accessed 23-September-2008].
- [51] U. v. Zadow. ibavg: A high-level Multimedia Platform. <http://www.libavg.de>, note = "[Online; accessed 23-September-2008]".

Filter type	No active blobs		Five active blobs	
CMU capture	3.351 ms	10.49%	3.048 ms	9.63%
Background removal	0.569 ms	1.78%	0.565 ms	1.78%
Simple highpass	4.751 ms	14.87%	4.788 ms	15.12%
Scaler	2.767 ms	8.66%	2.806 ms	8.86%
Barrel distortion correction	19.962 ms	62.49%	19.913 ms	62.90%
Rectify	0.544 ms	1.70%	0.538 ms	1.70%
Total filter time	31.944 ms	100%	31.658 ms	100%
Finding blobs	1.276 ms	99.61%	1.491 ms	94.97%
Tracking blobs	0.004 ms	0.31%	0.061 ms	3.89%
Dispatching events	0.001 ms	0.08%	0.018 ms	1.15%
Total tracker time	1.280 ms	100%	1.570 ms	100%
Total Touchlib time	33.225 ms		33.228 ms	

Table 2: Touchlib image processing and blob tracker latency results.

Projector type	Native resolution	Latency	Comments
3M DMS 700	1024×768 @ 60 Hz	100 ms	short throw
Canon LV-S1E	800×600 @ 60 Hz	16.67 ms or less	office projector
Casio XJ-S30	1024×768 @ 60 Hz	16.67 ms or less	office projector
Epson EMP-400W	1280×800 @ 60 Hz	-16.67 ms	short throw
NEC WT610	1024×768 @ 60 Hz	16.67 ms or less	short throw
Sanyo PLC-XL50	1024×768 @ 60 Hz	16.67 ms or less	short throw
Sharp PG-A10X	1024×768 @ 60 Hz	16.67 ms or less	office projector
Toshiba EW25	1280×800 @ 60 Hz	16.67 ms or less	short throw

Table 3: Digital projector latency results.

Used projector	3M DMS 700	3M DMS 700 (improved touchlib)	Sharp PG-A10X (improved touchlib)
Camera shutter and readout	20 ms	20 ms	20 ms
Firewire transfer	30 ms	30 ms	30 ms
Touchlib	33 ms	33 ms	33 ms
Touchlib ‘bug’	32 ms	0 ms	0 ms
Context switch	10 ms	10 ms	10 ms
Digital projector	100 ms	100 ms	0 ms
Total latency	225 ms	193 ms	93 ms

Table 4: Comparing the total system latency of different hardware and software combinations. Latency time is measured in milliseconds.