Building interactive multi-touch surfaces

Johannes Schöning, Jonathan Hook, Tom Bartindale, Dominik Schmidt, Patrick Oliver, Florian Echtler, Nima Motamedi, Peter Brandl

Abstract Multi-touch interaction with computationally enhanced surfaces has received considerable attention in recent years. Hardware implementations of multi-touch interaction such as Frustrated Total Internal Reflection (FTIR) and Diffused Illumination (DI) have allowed for the low cost development of surfaces. Although many of these technologies and associated applications have been presented in academic settings, the practicalities of building a high quality multi-touch enabled surface, both in terms of the software and hardware required, are not widely known. We draw upon our extensive experience as developers of multi-touch technology to provide practical advice in relation to building, and deploying applications upon, multi-touch surfaces. This includes technical details of the construction of optical multi-touch surfaces, including: infrared illumination, silicone compliant surfaces, projection screens, cameras, filters, and projectors, and an overview of existing software libraries for tracking.

1 Introduction

Multi-touch technology has opened up a wide range of opportunities for interaction design. Relatively simple and inexpensive hardware and software configurations allow the development of interfaces with expressive gestural control and fluid multi-user collaboration. The underlying technology has existed in different forms since the late 1970s and multiple patents [1, 2, 3, 4, 5] demonstrate how multi-touch surfaces can be constructed. However, it was Hans 2005 presentation [6] of a low-cost camera-based multi-touch sensing technique, based upon Frustrated Total Internal Reflection (FTIR), which truly highlighted the potential for multi-touch interaction in the development of the next generation of human computer interfaces.

Johannes Schöning
German Research Center for Artificial Intelligence (DFKI), Campus D3 2, Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany e-mail: johannes.schoening@dfki.de

Hans system was both cheap and easy to build, and was creatively applied to illustrate a range of novel interaction techniques. Indeed, his YouTube demonstration captured the imagination of experts and laymen alike, and as a result, we have seen an explosion of interest in multi-touch interaction. Hardware implementations of multi-touch interaction such as FTIR and DI have allowed for the low cost development of surfaces and enabled much research exploring the benefits of multi-touch interaction [7, 8, 9, 10, 11, 12, 13]. Unsurprisingly, multi-touch surfaces have also found their way into the Hollywoods futuristic visions of human-computer interaction (e.g. "James Bond – Quantum Of Solace" and "The Day the Earth Stood Still" [14]).

Although many the multi-touch technologies, and associated applications, have been presented in an academic context, the practicalities of building a high quality multi-touch enabled surface, both in terms of the software and hardware required, are still not widely known. In this chapter we draw upon our extensive experience as developers of multi-touch technology to provide practical advice as the development and deployment of such systems. This includes technical details of the construction of optical multi-touch surfaces, including: infrared illumination, silicone compliant surfaces, projection screens, cameras, filters, and projectors, and an overview of existing software libraries for tracking. Our goal is to enable researchers to embrace multi-touch by providing the basic knowledge required to "build your own" multi-touch surface. Many of the established technologies, such as resistance, capacitance, or surface wave-touch screens, require industrial fabrication facilities beyond those available even to the academic research. In contrast, we focus exclusively on optical approaches to multi-touch sensing which can be developed and integrated with an interactive application by a moderately competent hobbyist. Optical approaches to multi-touch use image processing to determine the location of interactions with the surface. Typically using infrared illumination, their simple set-up means they have the potential to be extremely robust. Although not described here, in addition to FTIR and DI there are a number of less widespread, but related, approaches including Laser Light Plane and Diffused Screen Illumination (see [15]). In addition there are other upcoming techniques and directions as presented at CHI 2009 [16, 17] also not covered in this chapter.

In this chapter we give step-by-step instructions as to how to build interactive multi-touch surfaces, discuss the pros and cons of the different approaches and try to help the reader avoid the traps, which a novice may fall into when developing their first surface. The arrival of large consumer multi-touch surfaces is eagerly anticipated. Similarly, display, projection and other technologies on which optical multi-touch systems depend continue to advance apace. As a result the practical 'shelf-life of our contribution would appear rather short. However, we see our account as more than simply a detailed documentation of a critical point in time for user interface software and technology. Whilst many of the details that we describe provide a context to current academic research in multi-touch (far more than the conventions of academic publishing normally allow) we see considerable value in human-computer interaction researchers continuing to develop their own underlying hardware. Indeed, multi-touch as a paradigm is still very much in its infancy.

The production of commercial systems has the potential to simultaneously promote the emerging status quo as to what a multi-touch interface is, and frustrate further development. So long as researchers shape and control the underlying hardware the opportunity for more fundamental innovation remains open.

2 Non-optical multi-touch approaches

Before describing optical multi-touch systems in more detail its appropriate to review alternative technologies. There is no fundamental characteristic that that optical approaches superior to the alternatives. Indeed, many of these alternatives have already found their way into consumer products, albeit in smaller interactive surfaces (e.g. mouse pads on laptops and touch-screens in phones). However, as already described, the principal drawback is that resistance, capacitance, or surface wave-touch screens, require industrial fabrication facilities.

2.1 Resistance-based Touch Surfaces

Resistance-based touch panels generally consist of two clear sheets coated with transparent conductive substances such as indium tin oxide [18]. These surfaces are separated by an insulating layer, typically tiny silicon dots. The front of the panel is often made of a flexible hard coated outer membrane while the back panel is typically 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 both horizontally and vertically (by the controller) to resolve the exact position of a the touch event. Such touch surfaces have the advantage of low power consumption, are used in mobile devices such as the Nintendo DS, mobile devices and digital cameras, and can be operated using fingers or a stylus. However, resistance-based technologies generally yield low clarity interactive surfaces (i.e. 75%–85%) and additional screen protection cannot be added without significantly impacting on their sensitivity. More detailed information about classical resistance based (multi-) touch surfaces can be found in [19].

2.2 Capacitance-based Touch Surfaces

Capacitance based (multi-) touch surfaces can be broadly subdivided into two classes depending on the underlying sensing mechanism: (1) Surface Capacitance; and (2) Projected Capacitance. Both technologies were originally developed for single touch interaction, and one advantage of capacitive touch surfaces over competing technologies is their high clarity; making capacitive touch surfaces very suitable for use where the display and touch sensitive surface are integrated (i.e. beyond simple touch pads). Capacitive touch screens are generally durable, reliable and can be operated by any conductive device and hence are not limited to finger based interaction. However, they are relatively expensive to manufacture and are therefore usually reserved for use in rugged environments such as in public displays and industrial applications. Although it is possible to manufacture capacitive multi-touch surfaces, typically the number of simultaneous touches is limited by firmware and/or by the design of the controller. Furthermore, accuracy decreases when performing touches with more than one object, although a number of capacitance-based technologies have been developed that overcome many of these restrictions in order to allow many simultaneous touches (e.g. MERLs DiamondTouch [7]).

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 again using indium tin oxide [18] 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 or electrons in the horizontal and vertical directions thereby setting up a uniform electric field across the conductive layer. As fingers (and other conductive objects) are also electrical devices capable of storing charge and supporting 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. Panels based on surface capacitive technology can provide a high positional accuracy.

2.2.2 Projected Capacitive Touch Surfaces

Of all the technologies we describe projected capacitive touch devices are the most expensive to produce. Their performance is also rather worse than competing technologies; however, they afford superb mechanical resilience. Projected capacitive surfaces can also be covered by a non-conductive material (up to a maximum thickness of approximately 20mm) without significantly impacting on their functionality. When used for (multi-) touch displays, as described by Rekimoto [11]) a very thin grid of microphone wires is installed between two protective glass layers. When touched, the capacitance between the finger and the sensor grid, and the touch location can be computed based on the electrical characteristics of the grid layer. The accuracy of projected capacitive technology is similar to that of surface capacitive technology although light transmission is often better since the wire grid can be constructed such that it is nearly transparent. The technology is also very 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, compared to surface capacitance technology, multiple simultaneous touches can be more easily interpreted.

2.3 *Surface Acoustic Wave Touch Surfaces (SAW)*

In surface acoustic wave surfaces 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 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.

3 Optical Based Touch Surfaces

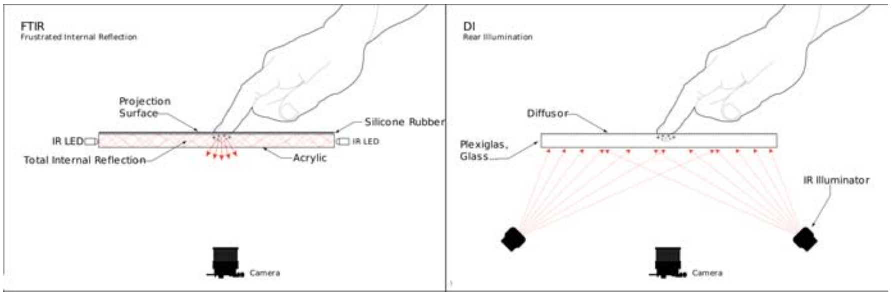

Fig. 1 General set-up of a FTIR system (left) and a DI system (right).

Optical approaches to multi-touch use image processing to determine the location and nature of interactions with the surface. These systems typically use infrared illumination, and due to their simple set-up have the potential to be very robust. Hans work in 2005 [6], which utilised the principle of FTIR in multi-touch interaction, can be seen as the turning point in both the interest and development of such optical systems. The FTIR approach is based on optical total internal reflection within an interactive surface. Electromagnetic waves transmitted inside a transparent surface are completely reflected if: (1) the inner material has a higher refractive index than the outer material; and (2) the angle of incidence at the boundary of the surface is sufficiently small. The most common FTIR configuration involves the use of a transparent acrylic pane into which infrared light is injected using strips of of

LEDs around its edges (see figure 1 (left)). When the user touches the acrylic, the light escapes and is reflected (due to its higher refractive index) and is reflected by the finger that is in contact with the surface. An infrared-sensitive camera aligned perpendicular to the surface can then clearly sense these reflections. A basic set of computer vision algorithms (see section 5) 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 multi-touch display (see figure 1 (left)). Diffuse Illumination (DI) systems have a similar configuration, with both a projector and an infrared sensitive camera placed behind the surface. However, for DI, infrared lighting is placed behind the projection surface; causing the area in front of the surface to be brightly lit in the infrared spectrum. As a result the camera is capable of detecting the infrared reflections of fingers and objects on, or in close proximity to, the surface (see figure 1 (right)). Touch detection exploits the fact that a projection surface (placed on the front of the surface) diffuses light, blurring objects at a distance. The main advantage of FTIR is that it allows very robust tracking of fingers, however, DI has the additional advantage that it allows easier tracking of physical objects, which can be identified either by their shape or through the use of fiducial markers [20] (easily recognizable markers usually in the form of a distinctive pattern) on the base of the objects. DI also has the potential to support hovering gestures, and any transparent surface (such as safety glass) can be placed between the projection screen and the user since sensing does not rely on surface contact.

4 Building Optical Multi-touch Surfaces: Step-by-step

In this section we divide the challenges faced in designing and building an optical multi-touch surface into those relating to the hardware and the software; our goal is to provide practical advice based on our own experiences of developing robust tabletop systems. The hardware of an optical multi-touch system comprises: infrared illumination sources, silicone compliant surfaces, projection screens (or LCD screens), cameras, filters, and projectors. In what follows we describe the desirable characteristics of each of these components and provide step-by-step advice on how they should be used. For a more detailed overview on the used material and possibilities to get them please refer to [21].

Step 1: Frame
The surfaces frame is the foundation of the whole system. For tablet-based surfaces it needs to be stable enough to support the wall covers, the surface, and all interior components; the projector being the heaviest of these. Crucially the rigidity of the structure is an important factor to consider when designing a frame as components will need to stay in place when the table is moved either for transportation or maintenance. Further considerations in re-

lation to the frame material include ease of assembly (as well as disassembly in case of tables that need to be transported) and weight, which is in particular important for larger structures. We suggest the use of an aluminium profile[1] system as it allows the realisation of a range of different structure and the profile itself is available in various thicknesses and lengths. Aluminium is easy to handle and craft, and is easily cut using a miter saw. In addition, a large range of flexible connection elements are available that allow the profile to be attached together in different configurations and for other components, such as projector or camera, to be easily mounted on the structure (see figure 2).

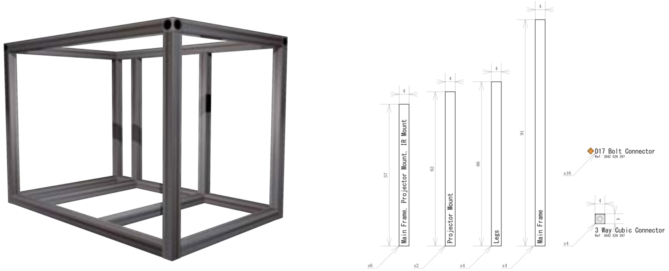

Fig. 2 Aluminium struts form a solid base structure for the interactive table. Due to the flexible connection system, components and walls can be easily mounted in any position. Twelve aluminium struts form the tables cuboidal main frame. Four more struts are inserted at the bottom and back to support the projector.

4.1 Displaying the GUI: Projectors and LCD Screens

An important issue, which must be considered when creating an optical multi-touch surface, is how the graphical user interface itself will be displayed upon the screen. The choice is usually between two technologies; using a digital projector, or using a modified LCD screen.

Step 2: Projector

A digital projector is used to project the interface image on the projection screen from behind the surface. Care must be taken to ensure that the projector has the appropriate resolution, throw and brightness, and the lag between input and output (in addition to sensing lag) is appropriate for the target application (see figure3).

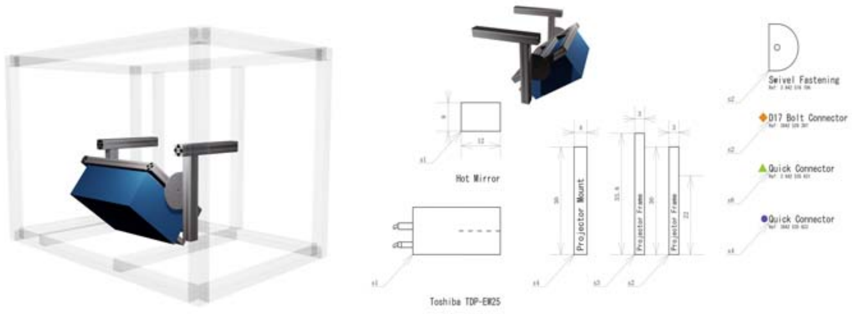

Fig. 3 With a single mirror, the 720p short-throw projector generates an image of about 107cm in diagonal at a table height of only 70cm. A hot mirror mounted in front of the projectors lens removes infrared emissions. The projector is held by two swivel fastenings attached to a supportive frame to allow for easy adjustments.

4.1.1 Projection based systems

Projection-based systems utilise a digital projector to display the interface upon a projection screen. As optical multi-touch surfaces commonly use transparent acrylic, it is possible to rear-project directly onto the interaction surface; preventing problem of occlusion that arises when projecting from the front. The use of a projector provides several advantages. Firstly, projection-based systems technologically simple to configure, with the projector being either pointed directly at the rear of the surface, or at the surface via a system of mirrors (see figure 3 for a range of configurations). Secondly, digital projectors are relatively cheap and can display a very large image. They are therefore an extremely cost affective solution for the creation of large interactive surfaces. Finally, as the image is projected onto the interaction surface, the use of appropriate optical filters means that a projection-based system will not interfere with the underlying sensing technology. Projection-based systems do however have a number of limitations. Firstly, digital projectors generally do not display an image at resolutions greater than XGA (1024 × 768). An image of this resolution, when projected over a large area, can lead to a low quality visual presentation, for example, one on which small or diagonally oriented text is difficult to read. High definition projectors are now commercially available, but are currently significantly more expensive than the mass consumer models. Another issue, which a developer employing a projector-based system will face, is *throw*; the distance between the projector and projection surface which is required to display an image of a specified size. When multi-touch surfaces are to be built in table forms or embedded in walls, this can be a significant challenge.

Projector throw: As described above, projector throw describes the distance from which the projector must be positioned from the projection surface for an image of a specified size to be displayed. The throw of a projector is an important element of its specification, and is typically given as two angles from which the size of the image (horizontally and vertically) at a given distance can be calculated. A standard digital projector (such as those for office use) will normally have a horizontal throw angle of no more than 35 degrees. To project an image upon a (1024×768)mm multi-touch table using a projector (which displays an image with an aspect ratio of $4:3$ with a throw angle of 35 degrees) the projector must be placed at least $1624mm$ away from the surface; this presents a problem as an average tabletop surface is only around $770mm$ above the ground. To address this we can simply use ultra short-throw projectors, which will project large images at very small distances; although again, the process can be prohibitive. A more cost affective alternative is to shorten the physical distance the projector must be placed from the surface by folding the projectors throw using a single, or series of, mirrors. When designing a system of mirrors a developer is likely to encounter two problems. Firstly angled mirrors can lead to distortion of the projected image (as seen in figure 3). This distortion can be corrected using the 'key-stone' function found on many projectors, although this can often lead to a lower quality image as pixels are stretched to counter the misshapen image. In most cases it is best to avoid such distortion by placing mirrors as close as possible to each other at right angles to, or 45 degree angles to, the projector (see figure 5). Secondly, a shadow image can result from light being reflected from not only the mirror itself, but also from the glass layer, which protects it. Such shadow images can be avoided by using a front surface mirror (which has no glass front) (see figure 4); these can be expensive and therefore the use of a mirror with a very thin front layer of glass is cheaper alternative, this will reduce, but not remove, the shadowing affect.

Brightness: In most applications the projector must be able to produce an image which is clear enough to be easily viewed by users even in a well lit room or in an environment where there is significant natural light. As a rule of thumb projectors need to have a brightness of at least 1500 lumens. In addition to brightness, the contrast ratio of a projector should be carefully considered, as more contrast allows for an image of similar clarity to be produced with a less bright projector (high contrast is particularly important for back-projected optical multi-touch surfaces.)

Resolution: The images produced by a standard XGA projector may not be of a high enough resolution to present certain content (such as text) when used for larger multi-touch surfaces. Therefore it is preferable when creating a multi-touch surface, which requires such content to (a) reduce the size of the surface; or (b) choose a projector with a greater resolution image; currently, the use of such high definition projectors is unlikely to be cost effective. A alternative solution, which has been demonstrated by Tuddenham and Robinson, is to use an array of lower resolution projectors to create a very high resolution image [22].

Lag: A final issue, which must be considered when choosing a projector for a multi-touch surface, is the projectors lag. Usually, an important quality of a multi-touch interface a sense of responsiveness to the user through the provision of an almost instant response to their interactions with the surface. Certain models of projectors can exhibit a slight delay in displaying the image upon the screen, when combined with the time delays introduced within the sensing pipeline this can contribute to an unresponsive interface. Schöning et al. [15] present an approach for measuring this time lag which can be used when choosing a projector for a multi-touch surface.

Step 3: Mirror
If a back projection is used, a mirror (of set of mirrors) may be used to fold the throw of the projector and thereby achieve a reduced table height. The use of a front surface mirrors to redirect the projectors light circumvents the problem of shadow images (see figure 4).

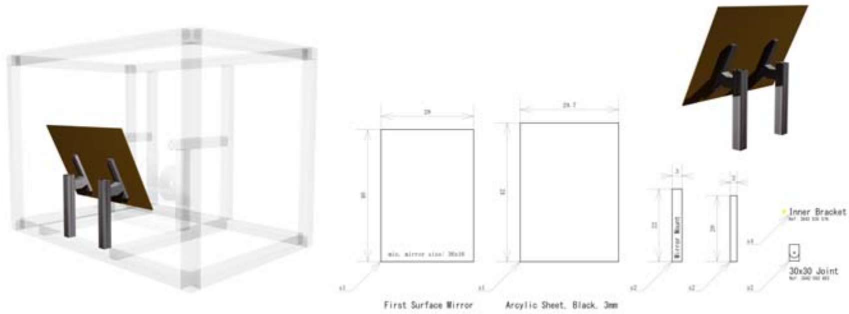

Fig. 4 Mirror: A front surface mirror redirects the projector's light without causing a shadow image.

4.1.2 LCD based systems

Using an LCD display instead of a projector has a number of advantages including the achievable DPI, sharpness, cost and size, but due to nature of current LCD technology, these benefits come at the expense of a reduced size of display region. Detailed information can be found in [23].

Choosing the right LCD: Almost any LCD screen (not Plasma or OLED) can be used in a multi-touch system, and although all LCD panels are in practice transparent to infrared light, two factors should be taken into account when selecting the device to use. Firstly, it is best to use a display with DVI or VGA input to avoid pixel information being lost when the video image quality is down-graded (for example with an LCD TV). Secondly, the aspect ratio should match to physical qualities of the system that you are intending to produce; widescreen LCD screens are now much more common and affordable.

Fitting the LCD: In order to use an LCD screen in a multi-touch display, the LCD glass panel inside the unit needs to be removed; it should consist of a thin sheet of black glass with a number of delicate circuit boards along one or two edges.

It is important to handle the panel with great care, so as not to scratch or chip this surface. Along with this panel, a control circuit, power supply and other boards may need to be removed and cables retained for later use. Sandwiched in-between layers in the display will be a number of optical sheets, including a Fresnel lens, diffusion and reflection layers; these should be saved these for re-fitting. Particular care must be taken when removing the backlights from the screen as they may contain harmful materials, and the control board may have a persistent high voltage charge. Once the LCD panel is removed, it can be fitted on the surface of a multi-touch display by supporting the surrounding edges firmly, and making sure all electronics are to the side and not in line of sight to the camera below.

Back Light: Unlike OLED displays, LCD displays only modulate visible light passing through them, so a source of white light is needed behind the display. Around 2500lm of visible light is needed; this light must be diffused as much as possible behind the display to avoid unattractive bright spots. This light source can be produced using the cold cathode tubes from the original display, traditional household lighting sources such as fluorescent tubes or LEDs. In all cases (except LEDs), it is important to be aware that most light sources produce ambient IR light, which may interfere with the tracking process. Some light sources may produce bright spots, so it is important to experiment with combinations of optical layers taken from the original LCD placed under the surface of the LCD to find the best diffuse light source without interfering with camera focus. Painting the inside of the unit white and removing large objects from the inside of the unit will also help to remove dark shadows from the displayed image, as when the screen is displaying white, it is effectively transparent. When using LEDs, the first thing to do is calculate how many will be needed in order to produce the desired luminosity (price also may be a consideration). Depending on the type of backlight in the original display, the perspex light box (with engraved or painted reflection tiles on one side) can be used against the underside of the LCD, experimentation will ascertain if the chosen camera setup will be able to focus through the backlight to the surface above.

Tracking Solutions: Due to the way in which LCDs are manufactured, IR light emitted from behind the display will reflect off its rear and not pass through in enough quantity to allow use of diffuse illumination, but any method of tracking with produces reflections of IR light from above the surface (e.g. FTIR) will produce the desired tracking effect. When calculating luminosity levels for infrared LEDs or lasers, be aware that the brighter the reflected image, the easier it will be for the camera to be over-exposed as the contrast ratio will be very small due to ambient radiation from the backlight. In order to track markers on an LCD based surface, the markers should be either active (contain a light source) or optically reflecting light from outside the unit through the display.

4.2 Tracking User Input: Cameras, Lenses and Filters

Optical multi-touch systems use cameras to track a user's interactions with the surface. In the case of FTIR and DI the camera captures the bright infrared light reflected from the users fingers as the make contact with the surface. In the case of tangible object-based interaction the camera tracks the fiducial markers on the underside of objects.

Step 4: Camera
User interactions are detected using a camera viewing the surface from below. The camera "sees" infrared light reflected by finger and objects touching the surface. The resulting image is then processed using computer vision software for touch detection. The projector and mirror are adjusted to produce an undistorted image, which fills the complete surface. While the camera is able to capture the whole surface area it does not suffer from interference from the projected image (see figure 5).

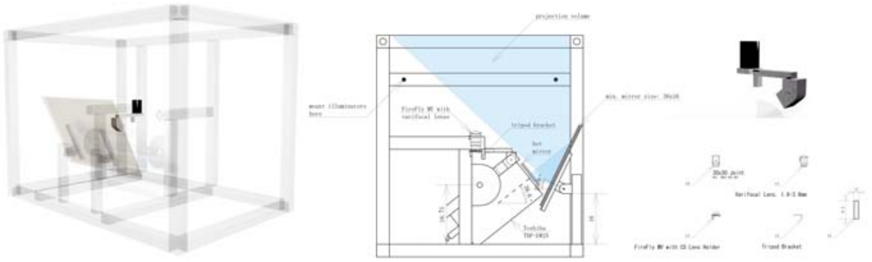

Fig. 5 Equipped with a wide-angle lens, a high-speed black and white camera can capture the whole surface at once. A band pass filter removes all visible light; only infrared light used to illuminate objects on the surface can pass.

4.2.1 Choosing the right Camera

Due to the important role the camera plays in an optical system, choice of the correct model is one of the most important design decisions. In the following sections we consider the impact of resolution, frame rate, sensor type, lenses and synchronisation of the camera.

Resolution: A camera must have a high enough resolution in order to allow it detect the small blobs of light which indicate that a user is making contact with the multi-touch surface. When designing a surface which will track touch interactions alone we have found that a camera with 640×480 pixel image will suffice for surfaces even as large as 1000×1000mm. When attempting to track tagged tangible objects however a much greater camera resolution is required to capture the detailed patterns of the fiducial markers. In this case we recommend a camera resolution of 1024×768 pixels (as a minimum). Camera resolution can impact on performance and responsiveness though, for as camera resolution is increases the time taken by tracking software to process the image also increases.

Frame rate: The frame rate of the camera directly affects the temporal sensing resolution of an optical multi-touch surface. Therefore if the surface is to feel responsive to the user then a camera with a frame rate of at least 30 frames per second (30HZ) is required. Increasing the frame rate of the camera will of course lead to a more responsive surface; for example, a camera which provides a frame rate of 60HZ will give a temporal sensing resolution similar to that of a mouse. Performance increases resulting from an increased camera frame rate may however be limited by the speed at which the tracking software can process the camera image.

Sensor type: As optical multi-touch surfaces are generally based around infrared illumination a camera with a sensor which can satisfactorily detect light in this range is crucial. Two decisions must be taken when selecting a camera, which can detect infrared light. Firstly, many digital video cameras are fitted with a filter designed to block infrared light. Consumer web-cams commonly have such filters, which are often difficult to remove, or are painted directly onto the lens. Therefore we recommend the choice of either an industrial grade camera as these generally have no such filters, or a camera with a filter that is easy to remove. Secondly, a sensor must be chosen which is sensitive to the bandwidth of infrared light emitted by the illumination source of the multi-touch surface (typically this will be $850nm$ or $880nm$). The data sheets of both the illumination source (such as the LED used) and the camera sensor will usually provide enough information to determine if the light emitted is within the range captured by that particular sensor.

Lenses: A camera must be selected with a field of view, which can capture the whole interaction surface from the distance allowed by the systems physical design. In many cases, such as when a camera is placed below a large interactive tabletop, this may be difficult to achieve with a standard lens; there are two solutions to this problem. Firstly the camera can be fitted with a replacement lens with a wider field of view; many consumer web-cams do not allow for lens replacement and hence an industrial grade camera may be required. It should be noted however that such 'wide-angle' lenses lead to a degree of fish-eye distortion of the image, which can prove problematic, especially when attempting to track fiducial markers. Secondly, as already described, a system of mirrors can be used to fold the distance required for the camera to view the required area.

Filters: Optical multi-touch systems can suffer from interference from environmental and other infrared light sources. The most problematic interference is created by the image displayed on the multi-touch surface (such as those from the projector). Also, ambient light from sources such as direct sunlight can be a significant problem. Such interference can be reduced by placing a filter over the camera which blocks unwanted light of specific wavelengths. Indeed, if a band-pass filter is used this will block out all light of a frequency outside a narrow band encompassing the wavelength of the illumination source. Suc dh band-pass filters are often expensive and so a cheaper alternative can be to use a low-pass filter which blocks only visible light from the camera. A visible light filter can be constructed cheaply from either a film negative, which has been exposed or from black acrylic. Although this will prevent inference from the projected image, problems may still arise if the surface is to be used in a room, which has bright natural or artificial light.

Camera Synchronization: An alternative (or complementary) approach to the reduction of interference from ambient light is to synchronise the camera shutter with a pulsed illumination source [24]. By operating the illumination source with short pulses, higher currents can be passed through the LEDs resulting in greater light emission. Also, by only opening the camera shutter for the short period during which the LEDs are illuminated, the amount of ambient infrared light, which reaches the sensor, can be drastically reduced and distinguished from the overall ambient level. Unlike a band-pass filter, this approach even reduces interference from ambient infrared light of the same wavelength as the illumination source. To utilise this approach however a camera, which can be controlled by a trigger signal, is required in addition to a configurable shutter speed.

4.3 Infrared Illumination

Step 5: Illumination
In the DI setup, infrared light sources emit light, which passes through the surface and is reflected by objects on top of (or even approach) the surface, thus making them visible to the camera. Using this principle, arbitrarily shaped objects and visual markers can be detected in addition to finger touches. In the FTIR set-up infrared light is injected into an acrylic surface around its edges. Fingers touching the surface will cause light to escape, resulting in bright and clearly visible touch points (see figure 6).

Both FTIR and DI require an infrared light source. Achieving the right infrared illumination can be challenging and requires a knowledge of both the different methods of illuminating a surface and different the types of IR LEDs (5mm, 3mm, SMD) that are available commercially. Almost all existing IR-based set-ups employ light-emitting diodes (LEDs) as light sources. Two commonly used types of IR LEDs are Osram SFH4250 (SMD) and Osram SFH485 (5 mm). Whether SMD devices or standard LEDs are more appropriate depends on a number of factors, for example, if the LEDs have to be mounted to the rim of an acrylic glass plate, this is easier to achieve with SMD sources, as it is possible to simply attach them to the rim with instant glue. After hardening, instant glue is chemically identical to acrylic glass creating a very strong, transparent bond. Mounting standard LEDs requires holes to

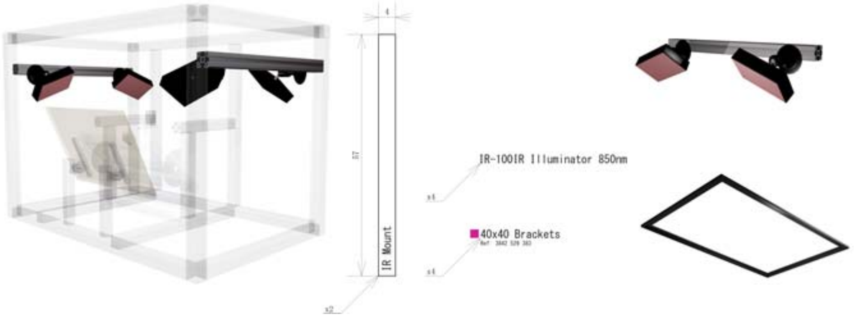

Fig. 6 Four infrared LED arrays are mounted in the table's corners, pointing downwards in order to diffuse the light and avoid overexposed areas.

be drilled into the material, which can be a time-consuming and error-prone process, and should be undertaken with care. One major problem for both FTIR and IR systems is their sensitivity to ambient IR light from the external environment. This 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 can be 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 level, which is possible during the short pulses, results in a much higher light output. The pulse duration and the following cool down period should be kept as close to the manufacturers 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. A second-level switching element is also necessary, to handle 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. Although more current will flow through the fuse than it is rated for, it is unlikely to blow during pulsed operation. Pulsing the LEDs significantly increases total light output, but this in itself does not produce enough contrast with ambient light levels. As already described, 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 pulses 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 environmental 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 down period between pulses, decreasing the contrast ratio. However, in the continuous mode, the brightness of the background is approximately 160 (when the LED is displayed with a maximum brightness – 255 in 8-bit mode), whereas in the pulsed mode, the background values are approximately 20, an eight-fold difference.

4.4 Surface: Material, Projection Screens and Compliant Surfaces

Step 6: Surface
In both FTIR and DI set-ups, a diffuser is required (in addition to the base acrylic sheet) to make the projectors image visible, i.e. serving as a back projection screen. Depending on the illumination technique applied, additional layers of different materials may have to be added (see figure 7).

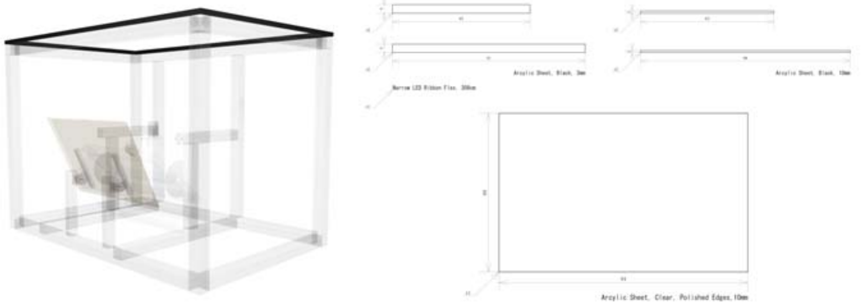

Fig. 7 A 1cm thick, transparent acrylic sheet forms the surface of the table in the DI set-up.

4.4.1 Surface materials for DI

DI requires a material that diffuses the light on the surface. This can be achieved by having the surface itself is as a diffuser or using a transparent surface with an additional diffuser. Plexiglas RP makes a good diffuse surface as in contrast to traditional Plexiglas, it has small micro-lenses embedded in the acrylic sheet that distribute the light evenly across the surface. The resulting projected image has no visible hotspot since the surface smoothes the light. Additionally, the gray surface allows for a good contrast with natural colours, and the material is scratch resistant and therefore well suited to direct touch interaction. As an alternative to Plexiglas RP a transparent surface material can be used combined with an additional diffuser. In this case a transparent acrylic plate is typically used as sturdy base layer (a common choice is a 5mm thick plate). In order to allow projections on the surface and to distribute

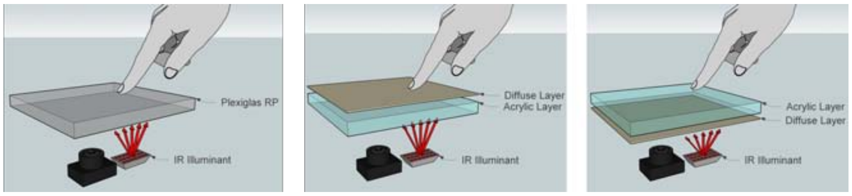

Fig. 8 The simplest DI setup uses the diffuse Plexiglas RP material (left). For Front DI setups, the diffuser is placed on top of the acrylic (middle). Rear DI needs the diffuser beneath the acrylic base layer (right).

the IR light, the additional diffuser is used. The diffuser can be a rear-projection foil or simply tracing paper. Any material used as a diffuser must allow enough IR light pass through so as to create a visible reflection on objects near the surface. The diffuser can be applied in front (front DI) or behind the acrylic (rear DI). If it is applied in front, the touch experience is generally more pleasant since the acrylic itself causes a high surface friction for dragging movements. However, the glossy backside of the acrylic result in hotspots due to the rear-mounted IR illuminators, which interfere with the vision tracking. If the diffuser is applied on the backside of the acrylic, this effect can be decreased, since the IR light is already diffused before it reaches the acrylic. Diffused Surface Illumination (DSI) is a variation on the standard DI setup. DSI uses a transparent acrylic called EndLighten that is a commercial lighting and presentation product. EndLighten has many embedded colourless diffuser particles, which distribute the IR light evenly across the surface. Instead of rear-mounted IR illuminators, DSI set-ups use an IR-LED frame similar to FTIR setups. The light from the LEDs is distributed and emitted uniformly over the entire surface. This results in a DI effect but with a FTIR light setup.

4.4.2 Surface materials for FTIR

Tables that use FTIR for tracking the users input are generally composed of a transparent acrylic plate augmented with a frame of IR-LEDs. The acrylic acts as a sturdy base layer that enables the FTIR effect. Additionally, the set-up needs a projection layer that is applied on top of the acrylic plate. Such an approach can negatively impact on its sensitivity and users have press hard on the surface in order to trigger the FTIR effect. Additionally when dragging a finger on the surface, such as in the performance of a motion gesture, friction may reduce the FTIR effect. As a result, many people use an additional layer (compliant surface layer) on top of the polycarbonate material to improve the sensitivity of the surface. These compliant surfaces are usually a soft and transparent material, which is placed between the polycarbonate sheet and a diffuse (projection screen) layer. Figure 9 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; the compliant surface layer intensifies this effect. Finding the correct

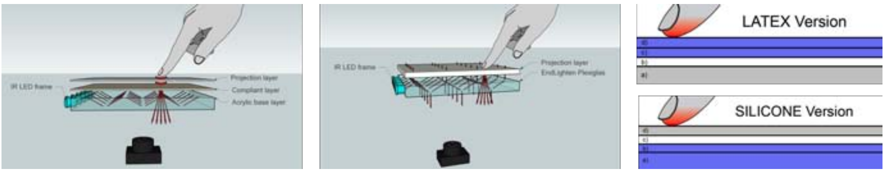

Fig. 9 The EndLighten Plexiglas distributes the IR light from the LED frame across the surface which creates a light emitting layer that can be used for DI tracking (left). The three layers needed to track the finger touches. The acrylic plate is covered with a compliant surface layer and a diffuse projection layer on top (middle). With the latex version, the projection (d) and the latex layer (c) must be combined; the gap (b) is between these two and the polycarbonate plate. In the silicone version, the gap (c) is between the projection surface (d) and the combined silicone (b) polycarbonate (a) layer (right).

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 or it sticks to the surface, constantly triggering the FTIR effect even after a finger has been removed. Very practical materials come in the form of SORTA-Clear$^{\text{TM}}$40 and ELASTOSIL®RT 601 silicone, both materials being relatively hard (Hardness Shore A ≥ 40), non tacky and very clear. Once hardened, both silicone layers can easily be removed from and re-attached to the polycarbonate surface. Using silicone as a compliant surface poses one problem however as the material comes as a gel which must be poured evenly over the surface. This can prove a difficult task. ELASTOSIL®RT 601 is less viscous and hence easier to pour, resulting in fewer bubbles in the vulcanized layer. As an alternative to silicone for the compliant layer, a thin layer of latex also works well. This also has an advantage over silicone layers as it does not have to be poured, reducing production time for the combined layer set-up significantly. Additionally latex is easier to handle, faster and cheaper to produce and more easily accessible as a mere 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 screens. However, in contrast to silicone, latex must be combined with the projection layer; with an air gap between the latex and the polycarbonate base plate, whereas in the silicone version we have exactly the opposite requirement. Figure 9 shows this difference between the latex and silicone layer construction. Depending on the compliant material (silicone or latex), it is possible to use different layers as the projection screen. The main requirements are to achieve a result that allows for an air gap between the correct two layers and the triggering of the FTIR effect. Rigid PVC is optimal for FTIR yielding a high contrast touch point. Comparable tracking can be achieved using tracing paper, but with a lower image quality and less robust characteristics. Materials such as Rosco translucent result in touch points that are either too dark, or showed permanent traces on the silicone. Other materials completely stuck to the silicone (HP backlit UV), are all considered not to be suitable for FTIR. For the latex version, HP Colorlucent Backlit UV is an effective option. HP Colorlucent Backlit UV foil was originally designed for use in back lighted signs. Similar to rear-projection screens it generates an even diffuse image without any hotspots from the projector, making it an ideal rear-projection surface. Because of its glossy backside, it cannot be used with the silicone version (as it sticks to the silicone). Rosco screens can also be combined with latex, since the latex sticks well to the screen.

5 Software

With the hardware in place, the next challenge is the selection and configuration of software components, with the goal being to set up a pipeline of image processing operators that transform a camera image into user interface events.

5.1 FTIR Tracking Pipeline

Figure 10 shows the canonical 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 [25]) finds bright regions in the pre-processed image; these are determined to be areas where a surface contact has been made. Post-processing involves finding corresponding touches in different camera frames (temporal correlation) and transforming the camera coordinates to screen coordinates.

5.2 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 appear blurred. reacTable [26] does this by adaptive thresholding based on the curvature of the luminance surface (see [27] for a detailed description of the algorithm). The multimedia platform *libavg*, [2] used in the *c-base MTC*, pioneered the use of a high-pass filter to achieve the same effect. Figure 10 shows images generated in a typical DI tracking pipeline. As can be seen, the image pipeline is split and the connected components algorithm is run 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 10 shows the result of this process: Fingers touching the surface have been identified and associated with hands.

---

[2] http://www.libavg.de/

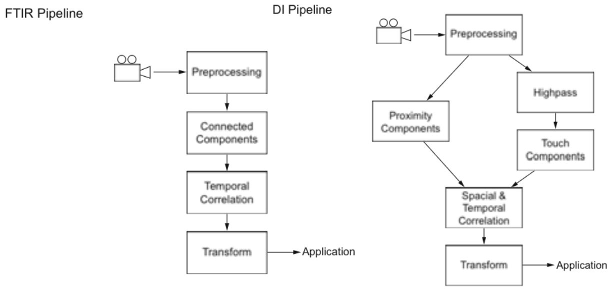

Fig. 10 FTIR Tracking Pipeline (left). DI Tracking Pipeline (right).

5.3 Interface Considerations

The tracking pipeline provides higher-level software layers with information about finger and hand positions. The TUIO protocol [28] uses Open Sound Control over UDP to transmit this information in a format, which can be interpreted easily by a wide variety of tools and languages. By default, Touchlib, and many other libraries, come with the functionality to send TUIO events over the popular OpenSound Control (OSC)[3] protocol. OSC libraries are available for many modern programming languages such as C#, Adobe Flash (Actionscript 3), Java, Max/DSP, Processing, Pure Data, Python and Visual Basic. When using Flash UDP packages have to be converted to TCP using the Flosc tool, which acts as a proxy. Other software packages which provide tracking and TUIO (or similar) output include TouchLib (and its successor, Community Core Vision), VVVV, OpenTouch, OpenFTIR, T-Labs multi-touch lib, libavg and libTISCH. For a more detailed overview please refer to [15].
Work is in progress to provide higher-level interfaces such as widget libraries (*libavg*, *NUI Suite Snowflake*[4], *libTISCH* [29]). *libavg* includes event processing that correlates touches to a hierarchy of on-screen widgets. *libTISCH* provides a hierarchy of layers such as tracking, calibration, interpretation of gestures and display of widgets [30]. 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 through window-system-like event processing. When an application uses the OSC protocol, it is only able to receive events containing properties of the detected blobs. It is for example not possible to adjust the settings of TouchLib from within 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.

---

[3] http://www.cnmat.berkeley.edu/OpenSoundControl/

[4] http://natural-ui.com/solutions/software.html

6 Lessons Learned and Outlook

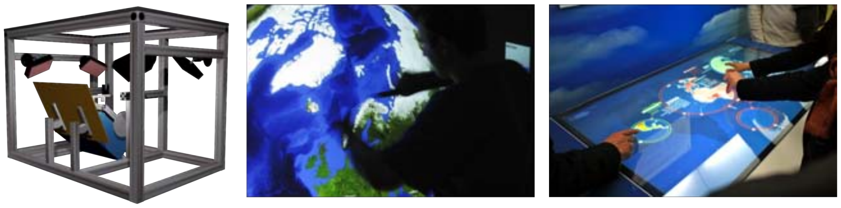

Fig. 11 A self-made interactive multi-touch table (left) and multi-touch applications on a vertical and horizontal multi-touch surface (middle, right). The application in the middle is described in more detail in [31]. The application on the right, is the GlobalData application by Archimedes Products. The application was part of GlobalData application was part of the Science Express Germany Exhibition.

The steps involved in building a high quality multi-touch enabled surface are not trivial. We hope that by sharing our knowledge and experience of developing multi-touch technologies with the wider community we can inspire the development of new tabletop interfaces, which embrace the possibilities, posed by this exciting interaction technology. Of course there are a lot of other helpful tutorials on the web that summarize knowledge on how to build multi-touch surface (one example is the Wiki book of the NUI group[5]). In this chapter we try to summarize all key information needed to build your own multi-touch surfaces and this chapter as a good starting point for multi-touch newbies. More advanced topics will be covered later in this book (e.g. bringing tangible interface onto multi-touch surfaces). Of course, despite the technologies described in this paper many fundamental questions for researchers still remain, including: What are the practical benefits of multi-touch systems over single-touch systems? What can graphics and interaction design practitioners "do" with multi-touch surfaces? Which applications is multi-touch input appropriate, viable and useful? Are there more than interaction possibilities than "just" rotating and scaling photos or zooming into maps? Is rotating a picture really a natural gesture? We hope that our description of the realities of building optical multi-touch surfaces will enable more people (experts and laymen alike) to engage in answering these questions and help then build more useful and interesting applications of interactive multi-touch surfaces. Multi-touch is probably here to stay, but as Bill Buxton said of the mouse: "Remember that it took 30 years between when the mouse was invented by Engelbart and English in 1965 to when it became ubiquitous". To speed this passage from invention to adoption we would like to encourage developers to design interfaces that help users forget the dominant WIMP paradigm of desktop computing, by producing designs that can only be operated using multi-touch gestures. More than this they should take their newly built interactive surfaces outside

---

[5] http://nuicode.com/projects/wiki-book

the lab and engage with users in the wild. Many of the most interesting and exciting observations as to the true utility of multi-touch has resulted from real-world observation of their use as in the City Wall project [32] or the multi-touch wall “Hightech Underground” [31]. As mentioned in the introduction, building interactive multi-touch surfaces and letting researchers shape and control the underlying hardware, gives the opportunity for more fundamental innovation.

Acknowledgements We would like to thank Florian Daiber, Otmar Hilliges, Markus Löchtefeld, Laurence Muller, Tim Roth, Daivd Smith, Urich von Zadow, David Hollmann & Antonio Kruüger for their help with the Multi-touch Bootcamp 2008 in Amsterdam in conjunction with IEEE Tabletops and for their comments and help (e.g. graphics and feedback) on this chapter.

References

1. Johnson RG (1972) #3,673,327: Touch actuable data input panel assembly. US Patent URL http://www.google.com/patents?vid=USPAT3673327
2. Kasday LR (1984) #4,484,179: Touch position sensitive surface. US Patent URL http://www.google.com/patents?vid=USPAT4484179
3. Mallos JB (1982) #4,346,376: Touch position sensitive surface. US Patent URL http://www.google.com/patents?vid=USPAT4346376
4. Mueller RE (1974) #3,846,826: Direct television drawing and image manipulation system. US Patent URL http://www.google.com/patents?vid=USPAT3846826
5. White RM (1987) #4,484,179: Tactile sensor employing a light conducting element and a resiliently deformable sheet. US Patent URL http://www.google.com/patents?vid=USPAT4484179
6. Han JY (2005) 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, ACM, New York, NY, USA, pp 115–118
7. Dietz P, Leigh D (2001) DiamondTouch: a multi-user touch technology. Proceedings of the 14th annual ACM symposium on User interface software and technology pp 219–226
8. Kin K, Agrawala M, DeRose T (2009) Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation. In: GI ’09: Proceedings of Graphics Interface 2009, Canadian Information Processing Society, Toronto, Ont., Canada, Canada, pp 119–124
9. Moscovich T (2006) Multi-touch interaction. In: CHI ’06: CHI ’06 extended abstracts on Human factors in computing systems, ACM, New York, NY, USA, pp 1775–1778
10. Moscovich T, Hughes JF (2008) Indirect mappings of multi-touch input using one and two hands. In: CHI ’08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, ACM, New York, NY, USA, pp 1275–1284
11. Rekimoto J (2002) Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In: CHI ’02: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, New York, NY, USA, pp 113–120
12. Schöning J, Hecht B, Raubal M, Krüger A, Marsh M, Rohs M (2008) 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, ACM, New York, NY, USA, pp 129–138
13. Valli A, Linari L (2008) Natural interaction sensitivetable. In: CHI ’08: CHI ’08 extended abstracts on Human factors in computing systems, ACM, New York, NY, USA, pp 2315–2318
14. Schöning J, Krüger A, Olivier P (2009) Multi-touch is dead, long live multi-touch. CHI 2009: Workshop on Multi-touch and Surface Computing

15. Schöning J, Brandl P, Daiber F, Echtler F, Hilliges O, Hook J, Löchtefeld M, Motamedi N, Muller L, Olivier P, Roth T, von Zadow U (2008) Multi-touch surfaces: A technical guide. Tech. rep., Technical University of Munich
16. Harrison C, Hudson SE (2009) Providing dynamically changeable physical buttons on a visual display. In: CHI '09: Proceedings of the 27th international conference on Human factors in computing systems, ACM, New York, NY, USA, pp 299–308
17. Rosenberg ID, Grau A, Hendee C, Awad N, Perlin K (2009) Impad: an inexpensive multi-touchpressure acquisition device. In: CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, ACM, New York, NY, USA, pp 3217–3222
18. Wikipedia (2009) Indium tin oxide — Wikipedia, The Free Encyclopedia. [Online; accessed 10-July-2009]
19. Downs R (2005) Using resistive touch screens for human/machine interface. Tech. rep., Texas Instruments Incorporated, Texas Instruments, Post Office Box 655303, Dallas, Texas 75265
20. Costanza E, Robinson J (2003) A region adjacency tree approach to the detection and design of fiducials. Vision, Video and Graphics (VVG pp 63–70
21. Schmidt D (2009) Design and Realization of an Interactive Multi-Touch Table. Technical Report, Lancaster University
22. Tuddenham P, Robinson P (2007) Distributed Tabletops: Supporting Remote and Mixed-Presence Tabletop Collaboration. In: Horizontal Interactive Human-Computer Systems, 2007. TABLETOP'07. Second Annual IEEE International Workshop on, pp 19–26
23. Motamedi N (2008) 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, ACM, New York, NY, USA, pp 3069–3074
24. Echtler F, Sielhorst T, Huber M, Klinker G (2009) A short guide to modulated light. In: TEI '09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction, ACM, New York, NY, USA, pp 393–396
25. Han Y, Wagner R (1990) An efficient and fast parallel-connected component algorithm. Journal of the ACM (JACM) 37(3):626–642
26. Kaltenbrunner M, Bencina R (2007) reacTIVision: a computer-vision framework for table-based tangible interaction. Proceedings of the 1st international conference on Tangible and embedded interaction pp 69–74
27. Costanza E, Shelley S, Robinson J (2003) Introducing audio d-touch: A tangible user interface for music composition and performance. Proc of the 6th Intl Conf on Digital Audio Effects (DAFX pp 63–70
28. Kaltenbrunner M, Bovermann T, Bencina R, Costanza E (2005) TUIO: A protocol for table-top tangible user interfaces. Proc of the The 6th Intl Workshop on Gesture in Human-Computer Interaction and Simulation
29. Echtler F (2008) TISCH: Tangible Interactive Surfaces for Collaboration between Humans. http://tisch.sourceforge.net/, [Online; accessed 23-September-2008]
30. Echtler F, Klinker G (2008) A multitouch software architecture. In: Proceedings of NordiCHI 2008, pp 463–466
31. Daiber F, Schöning J, Krüger A (2009) Whole body interaction with geospatial data. Smart Graphics pp 81–92
32. Peltonen P, Kurvinen E, Salovaara A, Jacucci G, Ilmonen T, Evans J, Oulasvirta A, Saarikko P (2008) It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In: CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, ACM, New York, NY, USA, pp 1285–1294