Literature review of Multi-Touch Interaction

Nicholas Irvine
Department of Electrical and Computer Engineering,
University of Auckland, New Zealand
Nirv002@ec.auckland.ac.nz

ABSTRACT

Multi-touch interaction in this case refers to human interaction with a computer where more than one finger can be used to provide input at a time [3]. The benefits of this are that the multi touch interaction is very natural, and it inherently provides support for simultaneous multi user input. There are several promising technologies that are being developed for multi-touch. The two main alternatives currently are capacitive sensing and Frustrated Total Internal Reflection (FTIR). Both of these technologies have been implemented successfully, and both have the possibility of being used in smaller consumer devices such as desktop workstation screens and mobile phones.

Multi-touch screens present new possibilities for interaction. Gesture systems based on multi-touch have been developed, as well as multi point systems. These systems have not been comprehensively tested on users.

Author Keywords
Multi-touch, touchscreen, multi-user interface.

ACM Classification Keywords
H.5.2 User Interfaces: Input devices and strategies

INTRODUCTION

While the idea of multi-touch has been around for many years, recent implementations offer a glimpse of how it could soon become much more common in day to day life. There is a range of technology being leveraged to support multi-touch, and research into the new forms of user interaction it can offer.

PROBLEMS

Many technologies have been used in small scale prototype implementations but few have made it into commercial success. Some that have are the Mitsubishi Diamond Touch, Microsoft Surface and the Apple iPhone.

These technologies revolve around several current promising techniques.

The usability of these screens has not been thoroughly empirically tested, but developers have anecdotal evidence to support their claims.

Multi-touch has applications in small group collaboration, military applications, modeling applications, and accessibility for people with disabilities

APPROACHES

There are many approaches being attempted to provide useful multi touch. None at the moment is perfect but they have their costs and benefits.

FRUSTRATED TOTAL INTERNAL REFLECTION

[3] Introduces the use of FTIR to create a multi-touch display which is scalable and cost efficient. FTIR has been used in fingerprint scanning for many years [3] and this paper suggests a scaled up version for the detection of fingers on a surface.

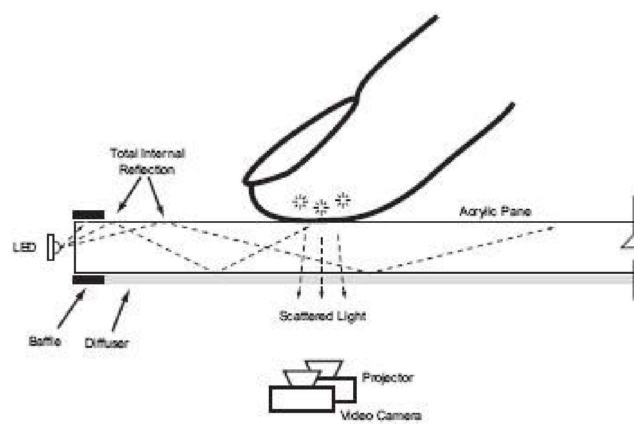

Figure1: FTIR [3]

FTIR works through a manipulation of the Total Internal Reflection of infrared light through a transparent surface. When a finger touches the surface, it causes a frustration of the internal reflection, and some light escapes, to be detected by a rear mounted digital camera. This occurs because the oils or sweat on the user's skin have a different refractive index than air, and so change the angle at which the light will be internally reflected. The image on the screen is projected by a projector mounted behind the interaction surface, next to the infrared camera.

The main benefit of using FTIR is the ability to exactly determine when a touch has taken place. This is because only a direct touch leads to the light being emitted and recorded by the camera.

A drawback of this system is that each touch is a discrete piece of information. It is not possible to discriminate between multiple touches of one person and different people touching the screen together. This makes the application of this technology in a multi user environment a problem, which is addressed in [2].

This FTIR system is suited to rear projection for the display of information, as a rear mounted camera is required to detect touch so a large depth behind the screen is needed. This makes the device not easily portable. These concerns have been partly addressed in [1].

This approach is scalable; it is limited only by the resolution of the infrared camera used, and the necessary brightness of the rear projection. A larger display requires a correspondingly brighter and more powerful projector.

This has been demonstrated by the author, who developed a multi-touch interaction wall 16 feet long by 3 feet high. Multiple users were able to walk up to the wall and interact with a range of applications.[ref to multi wall] This also demonstrates the inherent support for multiple users.

This same technology has also been developed into the Microsoft Surface.

HAND TRACKING

In this paper, [2] the authors have augmented the FTIR method to allow the touches of multiple users to be tracked and distinguished from each other. The augmentation also allows more accurate grouping of touches, so that multiple touches made by one person can be grouped to create arbitrarily complex gestures.

This augmentation is achieved by tracking the hands of individual users using an overhead camera. This also addresses a criticism raised in [3], relating to the effect of other light sources on the accuracy of the touch recognition. Infrared light from the surrounding environment may cause a touch to be detected in error. The addition of the camera image as a second reference makes this incorrect identification of touches less likely, as detected light emissions can be cross checked with the position of hands before registering as a touch.

This paper suggests two methods for discriminating between users' hands. The first is using skin colour segmentation. Previous research had shown that the intensity of skin colour is more important in distinguishing between people than the colour, so polarized lenses are used to remove the background image, and the intensity of the remaining image is used to distinguish users. The drawback is that the users' must wear short sleeved shirts to ensure their skin is visible to the camera.

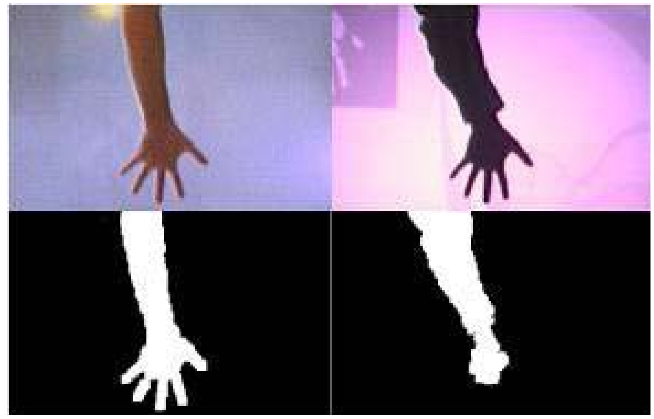

Figure 2 The images at left are the raw and processed hands using polarizing filters and skin detection. The images at right display RGB tracking without polarizing filters.[2]

The other method, used by the authors, is using RGB images from the overhead camera to generate an image of a user's hand by the shape of the shadowed areas. The drawback of this approach is that the image on the table surface cannot significantly change in brightness or intensity over the time of use, or the reference image that is removed from the camera image will no longer be valid.

For both methods, a 'finger end' and 'table end' of a user's arm are identified by the narrowing of the arm at the extremities. A user is identified partly by which side of the table they are on. This allows the assigning of a unique identifier to each user, so that their touches can be interpreted correctly.

A complex event generation method is presented, which allows user touches to be fired to many listeners, and be recorded in a user history.

The benefits of tracking the touches and associating them with a user are that touches can be recorded in history and can be linked together to create gestures. It also increases the scope for accurate multi user interaction.

MALLEABLE DISPLAYS

The authors of this paper [7] have taken the work done on multi-touch using FTIR and addressed several of the problems raised by Han in [3]. They have expanded upon this work by adding a malleable transparent silicone rubber overlay above the FTIR sensor to provide passive haptic feedback and pressure sensitivity.

One drawback of using FTIR as implemented by Han is that it relies on the sweat or oil on a user's skin to cause the frustration of the internal reflection and generate a touch event. In cold or dry environments, or for certain users, this can cause problems with detecting touches. The authors found in one situation approximately fifty percent of users encountered problems.

The addition of a soft flexible surface overlay mitigates this, because the overlay has a different enough refractive index that it will cause the frustration of the internal reflection whenever it is pushed against the detecting surface.

This also allows the use of other objects to generate input on the surface. Items such as paint brushes are used in an implementation of a painting program. The firm paint brush surface which would not by itself generate the FTIR does so when pressed onto the overlay.

The overlay was implemented in a 1cm thick version and a 1mm thick version. Both versions provided an improved touch detection rate, and pressure sensitivity.

Drawbacks of the thicker overlay included the need to use top down projection to display the interface, as the overlay was too cloudy, and a slight 'fuzzing' of images resulting in slightly less accurate shape recognition. The thick overlay however did provide improved pressure sensitivity.

Future work could involve using this thick malleable overlay to create real 3D terrain for mapping applications.

The current need for a large depth behind the screen and the use of rear projection could be replaced by using a standard LCD panel (which allows IR light to pass through it) as a display surface, and a mesh of IR sensitive photodiodes behind it to record the touch events.

TOUCH DETECTION WITH OVERHEAD CAMERAS

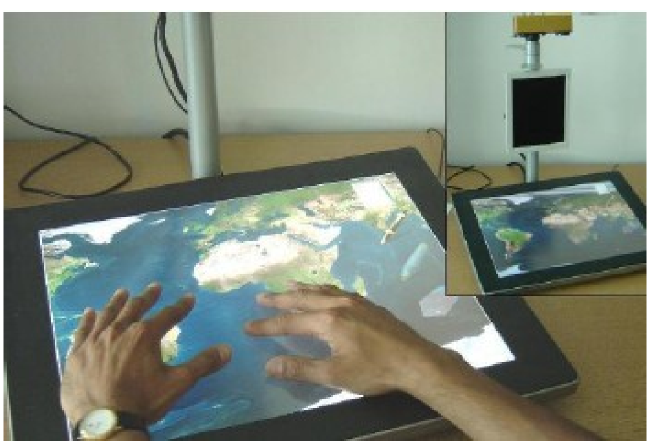

Figure 3: Set up for multi-touch using overhead cameras[1]

This paper [1] advocates a different approach to multi-touch interaction that does not use FTIR.

It uses 2 overhead cameras track the positions of user's fingers and to detect touches. The major problem with existing camera based systems is that they lack the ability to discern between a touch and a near touch. This places limits on the way a user can interact with the surface.

This paper provides a novel algorithm, developed by the authors to overcome this problem. This algorithm uses a geometric model of the finger and complex interpolation, which allows an accuracy of touch detection of around 98.48% [1].

The algorithm relies on 'machine learning' methods, where the algorithm is 'trained' by running it over many images of users hands on or near the surface.

The focus here is giving non multi-touch enabled surfaces or screens the appearance of multi-touch. Unlike in [2], the aim is not to support multi user multi-touch directly, but multi user support has been implemented by overlaying the image (captured by the overhead cameras) of one users hand onto the workspace of another, using degrees of transparency of the other users hand image to represent height. This means that the problem of determining which touch belongs to which user is circumvented.

The use of 2 overhead cameras means that any surface can be used to accept a user's touch. The authors have used this to provide multi-touch on a tablet PC.

MITSUBISHI DIAMOND TOUCH

Mitsubishi Electric Research Labs have developed a novel multi-touch table called the Diamond Touch[4]. It is designed for small group collaboration.

It has a mesh of antenna under the table surface. Each user has a receiver attached to them capacitively (not necessarily physically), often through the chair they sit upon. This is linked to the table through a USB connection. When a user touches the table the touch completes the circuit and sends a very small amount of charge through the user into their receiver. The receiver can determine what area of the table the user is touching by processing the signal. This allows unique identification of users, so that multiple users can be distinguished and combined gestures can be captured.

This implementation is top projected but could be made rear projected by using transparent electrodes and a transparent table surface.

One interesting property of this system is that there can be interference between users if they are sitting close together or physically touching. In this case the touches of either user can be interpreted as touches for both users, providing an easy and intuitive way of jointly selecting or indicating areas.

Another positive feature is that due to the matrix of antenna, often a user's hand will generate a capacitive signal in several neighboring antenna. Processing can be done to identify the midpoint of the antenna if the distance apart is small, indicating a specific touch, or the two aerials used to define an area if the distance between them is large. This allows the user a natural way of selecting a bounding box.

CAPACITIVE SENSING

This paper [5] uses capacitive sensing with a mesh of antenna to detect hand position as well as height above the surface.

A benefit of capacitive sensing is that it does not suffer from lighting related problems like FTIR, as touches are detected by electrical means. This does mean that it can suffer from electrical interference but this problem is overcome using a technique called lock-in amplifier.

A users hand can be detected when it is within 5-10cm of the surface. This allows the sensing of height and this can be used to create interesting interactions. This is an example:

Figure 4: A user 'picking up' an object on the screen [5]

A user can 'pick up' an object on the screen by drawing their finger tips close together over the object and raising their hand off the interaction surface.

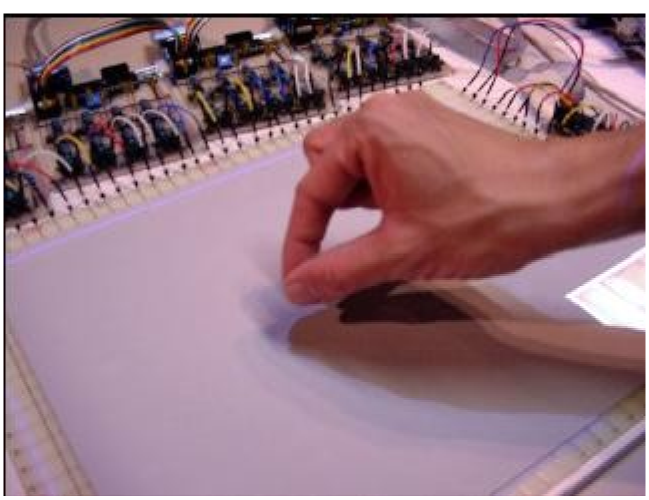

Figure 5: The result of the user lifting their hand off the screen after 'picking up' an object [5]

They can then move their hand around the table surface and 'drop' the object back onto the table surface.

In addition, the system can independently track many hands at once, which increases the scope of interaction by naturally supporting multiple user's at once.

It is constructed using a mesh of electrodes under a surface, similar to the Diamond Touch. These electrodes can be as simple as copper wires, or as complex as Indium-Tin Oxide (ITO) or a conductive polymer. These would allow flexible or transparent displays. Because modern flat screens use transparent electrodes in their design they could be very easily integrated with this technology. An example of this would be the Apple iPhone which uses capacitive sensing.

The resolution of the sensing can be improved by decreasing the distance between the electrodes. Electrodes spaced 10cm apart provided a resolution accurate to around 1cm.

INTERACTION TECHNIQUES

The development of technology to support multi-touch screens has been matched with the development of multi-touch and gesture based methods of interaction.

There has been work done on gesture recognition for those with physical disabilities, as often gestures normally require a full range of hand function and are difficult to perform for those with limited range in their fingers or wrists. People with disabilities such as these are often still able to use parts of their palm or the lower side of the hand to gesture.

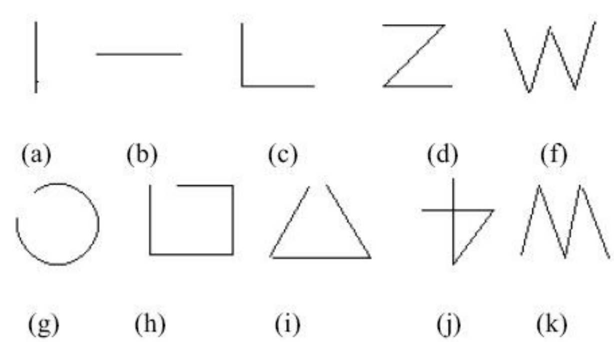

Figure 6: An example set of gestures that was developed [6]

These gestures are command based, and gesture 'a' for example could be interpreted to mean go up a level in a directory structure, and 'b' to go forward in a web page. These gestures can be easily performed by users with limited hand function and this was shown through an experiment.

This is an example of using objects to interact with a capacitive touch screen. The block on the screen in Figure 7 itself does not trigger a touch event, as it is not conducting electricity. When the user touches the block a capacitive connection is created and the object registers as a touch. Objects can be uniquely identified and registered with the system, through a kind of 'barcode' system, allowing things such as using objects as commands. A particular object could be moved on the screen, indicating that data should be transferred from one place to another.

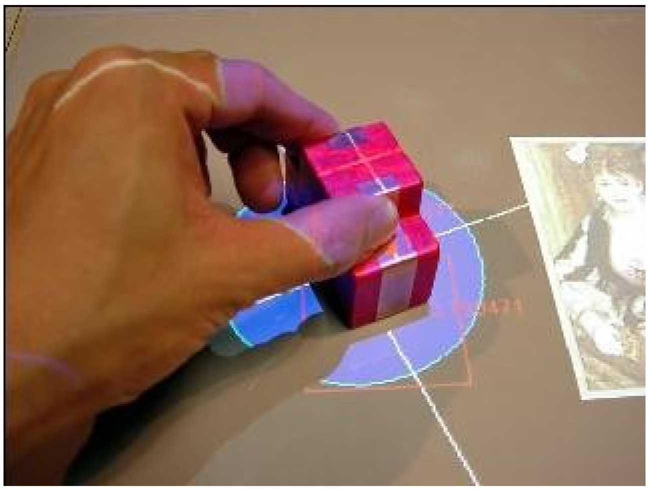

Figure 7: A user interacting with a capacitive screen using a block[5]

The mouse is a common form of input device. A problem with the mouse is that it does not represent how we manipulate things in real life. We often touch multiple points on an object's surface in order to manipulate it such as rotating it. This mean that in using a mouse we must reduce some tasks into simpler steps which may slow us down. Multi-touch allows a more subtle form of interaction that is more natural.

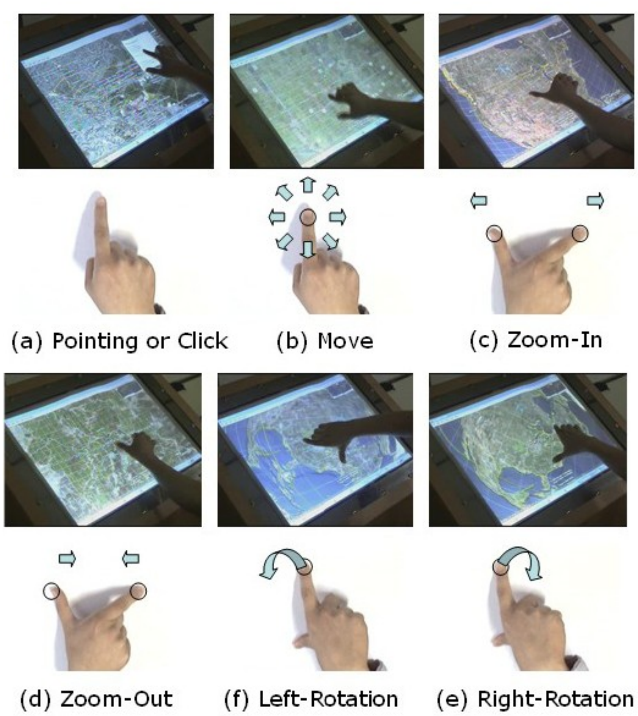

Figure 8: One handed gestures [6]

A common form of interaction that these authors present as being made possible by their work is the idea of a two point operation. This can be for example moving finger tips apart on a screen to zoom in, or move fingers apart to zoom out. Two fingers can be used in a twisting motion to rotate objects. This is advocated in [8] for a photograph viewing application.

In [6] a gesture library is developed for interaction with Google Earth. There were two sets of gestures, one for use with one hand and the other for two hands. These gestures were not tested on users.

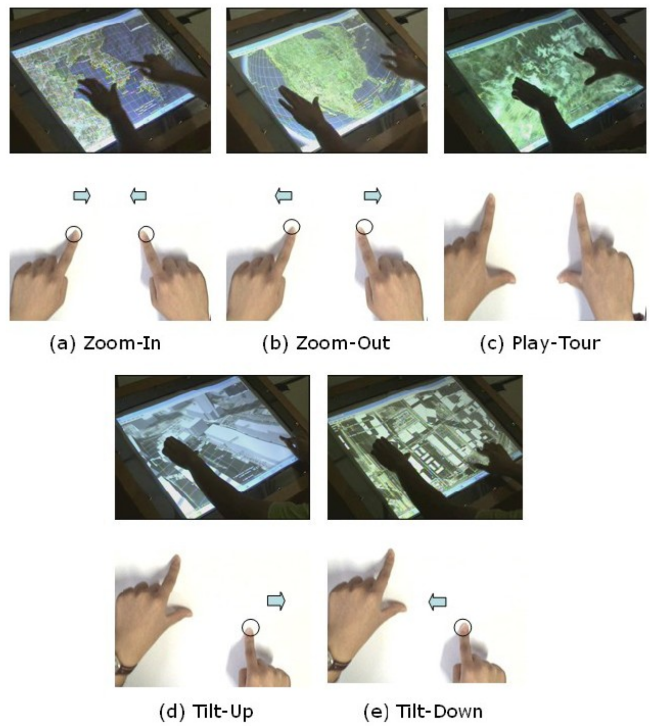

Figure 9: Two handed gestures[6]

SUMMARY

There are several promising technologies that are being developed for multi-touch. The two main alternatives are capacitive sensing and FTIR. Both of these technologies have been implemented successfully, and both have the possibility of being used in smaller consumer devices such as desktop workstation screens and mobile phones. The Microsoft Surface and the Apple iPhone are examples of successful commercialization of these technologies. Multi-touch screens present new possibilities for interaction. Gesture systems based on multi-touch have been developed, as well as multi point systems.

FUTURE WORK

The future work in this topic has been identified by the authors of many papers researched. There is a need to develop these prototype implementations further and integrate them with existing technology such as mobile phones or laptops. The possibility of making the sensing antenna transparent and removing the need for rear projection when using FTIR opens up scope for successful commercializations. Current problems with light and electrical interference need to be overcome to increase the robustness. The gesture systems being developed need to be tested on users and best practice developed so that future technologies can support reasonable gesture systems..

REFERENCES

1. Agarwal, A. Izadi, S. Chandraker, M. Blake, A (2007), High Precision Multi-touch Sensing on Surfaces using Overhead Cameras, *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems 2007*,197-200. Retrieved March 26, 2008, from IEEEXplore database. http://ieeexplore.ieee.org.ezproxy.auckland.ac.nz/search/srchabstract.jsp?arnumber=4384130&isnumber=4384097&punumber=4384096&k2dockey=4384130@ieeecnfs&query=%28%28izadi%29%3Cin%3Eau+%29&pos=11&access=no

2. Dohse, K.C. Dohse, T. Parkhurst, D. Still, J. (2008). Enhancing Multi-user Interaction with Multi-touch Tabletop Displays using Hand Tracking, *First International Conference on Advances in Computer-Human Interaction 2008*,297-302. Retrieved March 26, 2008, from IEEEXplore database. http://ieeexplore.ieee.org.ezproxy.auckland.ac.nz/search/srchabstract.jsp?arnumber=4455997&isnumber=4455943&punumber=4455942&k2dockey=4455997@ieeecnfs&query=%28%28dohse%29%3Cin%3Eau+%29&pos=0&access=no

3. Han, J. (2005, October 23-26). Low-cost Multi-Touch sensing through Frustrated Total Internal Reflection, *18th annual ACM symposium on User interface software and technology*,115-118. Retrieved March 26, 2008, from ACM database. http://doi.acm.org.ezproxy.auckland.ac.nz/10.1145/1095034.1095054

4. Dietz, P. Leigh, D. (2002). DiamondTouch: A Multi-User Touch Technology, *14th annual ACM symposium on User interface software and technology*, 219-226. Retrieved April 23,2008, from ACM database.

   http://doi.acm.org.ezproxy.auckland.ac.nz/10.1145/502348.502389

5. Rekimoto, J. (2002). SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces, *SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves,* 113-120. Retrieved April 22, 2008, from ACM database

   http://doi.acm.org.ezproxy.auckland.ac.nz/10.1145/502348.502389

6. Kim, J. Park, J. Kim, H, Lee, C. (2007) HCI(Human Computer Interaction) Using Multi-touch Tabletop Display, *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing 2007*, 391-394.

   Retrieved April 27, 2008, from IEEEXplore database

   http://ieeexplore.ieee.org.ezproxy.auckland.ac.nz/search/srchabstract.jsp?arnumber=4313256&isnumber=4313156&punumber=4313155&k2dockey=4313256@ieeecnfs&query=%28%28jangwoon%29%3Cin%3Eau+%29&pos=0&access=no

7. Smith, D. Graham, N. Holman, D. Borchers, Jan. (2007) Low-cost Malleable Surfaces with Multi-Touch Pressure Sensitivity, *Second Annual IEEE*

International Workshop on Horizontal Interactive Human-Computer Systems, 2007, 205-208

Retrieved 26 April 2008 from ACM database, http://doi.acm.org.ezproxy.auckland.ac.nz/10.1109/TABLETOP.2007.1

8. Han, J. (2006) Multi-Touch Interaction Wall, *ACM SIGGRAPH 2006 Emerging technologies, International Conference on Computer Graphics and Interactive Techniques.*