Human-Computer Interaction on Tabletops

Andreas M. Kunz
ICVR – Innovation Center Virtual Reality
ETH Zurich
Zurich, Switzerland
kunz@iwf.mavt.ethz.ch

***Abstract*— The most common working situation is standing or sitting at a table, and performing daily business work. Although this situation is very intuitive to the user, a computer support can hardly be found in this area, because of missing or inadequate human-computer interfaces that do not meet the user's requirements. At the ICVR, research in this field is done in a multi-disciplinary team, integrating expertise from electrical and mechanical engineering, computer science, and industrial design. The following paper shows some results of this interdisciplinary research and gives and outlook on future research.**

Keywords-HCI; tabletop interaction; tracking; multi-user

I. MOTIVATION

In private and office work, we are used to work at a table. This table is crowded with various things, which will be used for interaction, such as pen, paper, ruler, ink dwell, eraser, notepad, etc. Most of these devices show their intrinsic functionality by their well-known shape. Even more, we are also used to interact with fingers, e.g. when moving around a piece of paper. In summary, tabletop interaction in the 'analog' world contains multiple possibilities how the objects on the table can be modified or rearranged. For performing the given task, humans use color and shape detection of objects, haptic perception, and sometimes also acoustic hints to detect objects or an action. Moreover, human beings are sensitive against temporal delays (latencies), and have a large knowledge in realistic correlations between visual, haptic, and acoustic cues.

Taking this complex interaction procedure into account, it is obvious that various disciplines and expertise are required to realize an IT-supported (Information Technology, IT) tabletop environment, which supports the user's natural behavior when performing a task. The interactive devices (Tangible User Interfaces, TUI) need to be tracked, identified, and also possible states must be read out. To control the digital content on the table using TUIs, they need to be dynamically or statically assigned to a virtual object, while the system's reaction should not have any delay. Further, there is the requirement that multiple devices should be simultaneously useable (often called multi-user capability) in order to allow bimanual (single user) or multi-manual (multi-user) interaction.

Since interacting on a tabletop significantly differs from the interaction on a vertical whiteboard, existing technologies cannot be reused. When writing onto a whiteboard for example, the user touches it only in one point (the tip of the pen). When writing on a tabletop instead, the user will touch the table at several positions, i.e. with the tip of the pen, with the palm of the hand, with the forearm, and even the non-dominant hand will rest on the screen. Thus, the IT-system needs to distinguish between unwanted interaction and interactions done on purpose. The same holds true for other objects on the table. While TUIs are meant for interaction and should be recognized by the system, other objects are not used as TUIs and should be ignored. Realizing tabletop human-computer interfaces also gives problems regarding imaging. When displaying information, a front- or a back-projection is used. Front-projections suffer from shadow-casting during an interaction, while back-projections usually do not allow for ergonomic sitting at the table. Thus, a flat screen would be preferable, which was for a long time too expensive or had an image diagonal that was too small and thus unsuitable for teamwork.

II. TOWARDS AN IDEAL TABLETOP HUMAN-COMPUTER INTERFACE

The following paragraphs introduce systems that were developed by the ICVR research group since 1995 (see also http://www.icvr.ethz.ch). In order to realize such complex systems and to adapt them to the users' needs, it is required to have various disciplines within the research group. Thus, the group integrates expertise from mechanical and electrical engineering, process engineering, computer science, industrial design, and even architecture. Bringing together these different disciplines also allows addressing various application fields, e.g. from mechanical engineering or architecture. It also brings multiple – very often contradictory – views onto one given problem. This stimulates the discussion and leads to a very convincing system in the end.

From the very beginning, the mission of the ICVR was to develop IT-supported systems that support all business processes in product development and in the field of the digital factory. This starts from the idea generation with writing and sketching on interactive surfaces), and ends in planning of complete manufacturing lines. In all cases, the intuitive and simultaneous interaction with the system was emphasized. In the early years, the group focused on VR-systems only, developing well-known systems like the "blue-c" cave setup. From this, the group moved towards collaboration setups, which could be integrated easily in daily business processes.

The following paragraphs give examples of the outcomes in the past decade, in particular in the field of tabletop interaction. While on the first glance these systems seem to require only a 2D-tracking, they turned out to become much more complex

ISSN: 2236-3297

when also considering multiple different devices, which also could have multiple working conditions, and which also should take in to account the z-position (above the tabletop).

A. *The BUILD-IT system*

In 1998, BUILD-IT [1] was introduced as one of the first multi-user systems. It uses a table, on which the virtual objects, e.g. machines for layout planning, are front-projected. The users can sit at the table to interact with the system. Passive so-called bricks are used as TUIs. A brick can also be used to control a camera, and its view is then shown on a vertical screen. For detecting the bricks, IR-light is projected onto the table, which is only reflected by a special foil on the bricks. The reflected IR-light is detected by a camera, which is equipped with an IR-pass filter. Thus, the camera only sees the reflecting bricks, but not the other objects on the table, which allows filtering TUIs from other objects. Since the bricks are rectangular, the tracking system can detect x- and y-position, as well as a rotation around the z-axis. To perform an interaction, the brick is dynamically assigned to a virtual object being displayed on the table. To release this assignment, the brick is covered by the hand, which blocks the optical tracking system, and then taken away (much like the grasping gesture).

Since all bricks are identical, it is not possible to distinguish them. They also do not have any special states (e.g. 'right mouse button'). Since picking and placing of objects is a coarse manual interaction, shadow-casting as well as latency did not proof to be any problem. The system is not capable of tracking finger-touch, and thus users can make natural gestures to explain topics without irritating the system.

Figure 1. The BUILD-IT system

B. *CRiON Table*

In 2002, a workplace to support teamwork [2] was introduced. It was designed for writing and sketching on a white surface. To avoid shadow-casting, a back-projection was used, and to distinguish between pens of different color, they were equipped with different light emitting diodes (LEDs).

Since the pen's LED worked in the visible spectrum, it had to be brighter than the back-projected image. This allows the camera underneath the table to see four different color spots much brighter than the projected image. Thus, the system could track and distinguish four different pens simultaneously. Beside the pens, no other interaction tools were provided, since the system was limited by the number of available LED colors. Only four basic colors red, green, blue, and white did work for defining the pens' ID. Since the camera's gain was reduced so that only the bright LED could be detected, no finger touch was possible.

Figure 2. The CRiON table

C. *InfrActables*

For precise interactions such as writing or sketching, it is crucial to avoid shadow-casting by using a back-projection. However, using visible light for sketching and identifying devices was not always stable, in particular in the projection hotspot. To overcome this problem, IR light was used instead., This is invisible to the human eye and thus any irritations were avoided. However, since only one certain wavelength of the infrared was used, a device distinction by colors, i.e. by different wavelength was not possible anymore. Thus, the InfrActables [3], [4] uses a 5-bit word to signalize ID and state. The principle setup and its realization are shown in Figure 3.

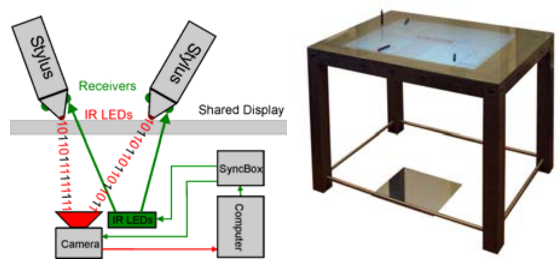

Figure 3. Principle setup of InfrActables and its realization

Three out of five bits are used for the ID, while two are used for the state. Since '000' for the ID and '00' for the state were not used to have at least two times a '1' for a more stable tracking, six different devices with three states each could be tracked and distinguished. The system synchronization triggers the camera and emits an IR-flash to trigger the devices via their IR-receiver. Now, five camera frames are acquired to achieve position, ID, and state (see Figure 3.).

InfrActables did not only realize pens of different colors, but also ruler, ink dwell, notepad, eraser, as well as a caliper option with the ruler (see Figure 4.). These devices were used to support digital brainstorming sessions.

Figure 4. Different devices in the InterActables system

For realizing these devices, also orientation and distance measurement was required. This was realized by using more IR-LEDs in such devices, which span an asymmetric shape that allows detecting the orientation unequivocally.

ISSN: 2236-3297

Using a standard camera for image acquisition, the refresh rate drops down to 12 Hz for a 60 Hz camera. This refresh rate is at the lower limit and results in a noticeable delay between the user's action and the system's response. To overcome this problem, a high-speed camera was used. Together with an optical bitcode using start and stop bits, it guaranteed a robust and fast identification and tracking [5]. However, once the distance of the devices' IR-blobs is below the camera's resolution, they cannot be distinguished anymore.

D. MightyTrace

So far, all systems relied on projections, resulting in shadow-casting in case of front-projection, and in non-ergonomic sitting (if any) in case of back-projection. Thus, MightTrace [6] tries to overcome theses problems associated with front- and back-projection. In the system, a standard LC-display is used. This allows an ergonomic sitting at the table, while interacting on the table, using the same type of devices as in InfrActables. Since there is no light-path available anymore for the camera, the camera is replaced by an array of IR-sensors, which is placed behind the LC-matrix (see Figure 5.).

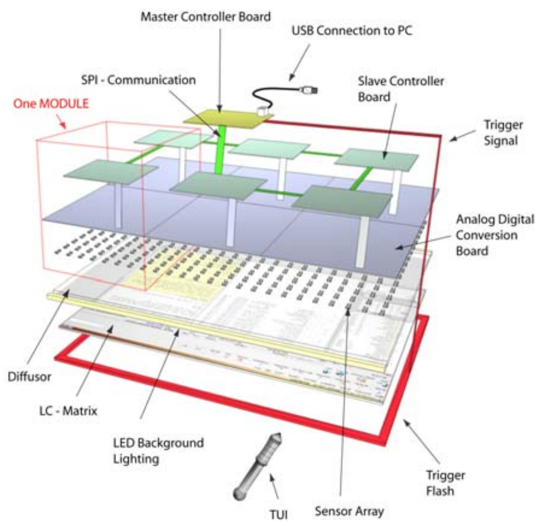

Figure 5. Principle setup of the MightyTrace

Since the LC-matrix is transparent to IR-light, the sensors can detect the IR-light sources in the active TUIs. Electronics for data pre-processing is also integrated in the LC-screen. Since the IR-sensors have a high sensitivity, they can sense signals at high refresh rate of around 2 kHz, while a standard CCD-camera only reaches e.g. 60 Hz. Now, a special transfer protocol can be used (see Figure 6.).

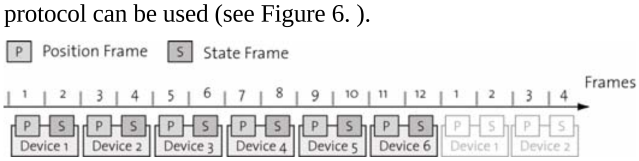

Figure 6. A complete frameset of 6 devices with their specific timeslots

Since all devices are sequentially interrogated, they can be unambiguously identified by the frame numbers in the temporal sequence. The second frame for each device is to signalize a possible state, i.e. two different states are possible (1 bit). The device's position can be determined by an interpolation of the sensors' signals, by which a resolution of 2 mm can be achieved. The time-sequential interrogation of the devices also completely avoids the garbling effect if two devices come close to each other. MightyTrace is capable of detecting and tracking active TUIs, while other objects and the fingers cannot be detected.

E. FLATIR

Although the LC-matrix is transparent to infrared light, there is certain absorption of the IR-light, which is negligible as long as the IR-light source is bright enough. However, it is not clear whether the sensors are sensitive enough to detect passive devices by the reflection of IR-light, as it would occur when tracking fingers on the LC-screen. Thus, FLATIR [7] was introduced, which uses Frustrated Total Internal Reflection (FTIR) to detect finger touch. IR-light is coupled into an acrylic layer and can be coupled out if the optical impedance changes, e.g. at the position where a finger touches the surface. Since the amount of light coupled out is very small, the sensors underneath the LC-matrix have to be placed closer to each to reach reliable touch detection and tracking (see Figure 7.).

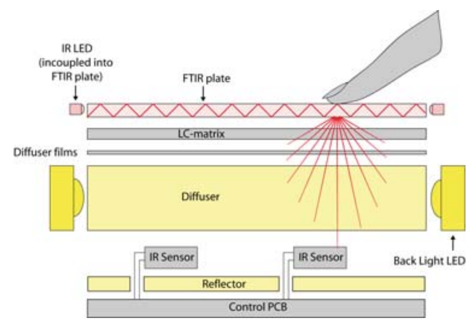

Figure 7. IR-sensors in the LC-screen

Now, finger touches on the LC-screen can be detected, using the same hardware as in MightyTrace. However, studies showed a shadowing effect, meaning that IR-light which is coupled out is missing at another position for detection. Further, the sensor's gain needs to be increased in such a way that any active device would drive the sensor into saturation, which significantly reduces the tracking accuracy.

F. TNT

In the TNT prototype [8], an attempt was made to combine finger touch and TUIs on the same electronics inside the LC-screen. For doing so, the gain for the IR-sensors has to be adjusted correspondingly to TUI interaction and finger touch by a modified time-sequential protocol shown in Figure 8.

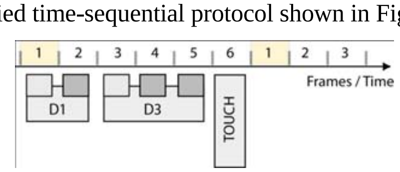

Figure 8. Modified sequential protocol [9]

ISSN: 2236-3297

The protocol offers an additional timeslot, in which only passive objects could 'respond' upon an IR-flash. For this timeslot, the sensitivity of the IR-sensors is increased to allow a good detection of the passive objects. In the next moment, the TUIs are triggered, the sensors' gain is reduced, and now only the strong IR-signal from the active TUIs can be detected. Due to the high update rate of the sensors, there is no noticeable delay by the sequential interrogation of active devices and finger touch. Even multiple finger touches can be detected.

G. *DigiSketch*

So far, the working environment is completely digital. However, users often want to integrate paper-based documents into this digital working environment, or at least use the same interaction device for writing on the paper and on the screen. Within DigiSketch [10], an attempt was made to use the commercial Anoto technology for this. The working principle of the Anoto pen relies on the reflection of IR-light on white paper, only partly absorbed by a non-repetitive black pattern being printed on the paper. A camera and a computer-vision system are integrated in a pen, which detects the absolute position on the paper and transfers this position to the computer via a Bluetooth connection. However, this principle does not work on LC-screens, since they are transparent to infrared.

In order to overcome this problem, an additional IR-reflecting layer together with the Anoto pattern was integrated in the screen. Although these additional layers also slightly affect the visible spectrum, the displayed artifacts are still visible with sufficient quality (see Figure 9.).

Figure 9. Using the Anoto pen on a LC-display

Since each pen carries its own tracking system, the amount of possible TUIs is theoretically unlimited. However, the Bluetooth data transfer implies a bottleneck, which currently limits the amount of possible devices to 6. In addition, touch is not possible with the Anoto technology, but could be easily added e.g. by an additional capacitive overlay.

III. SUMMARY AND OUTLOOK

This paper gave an overview of how to realize a human-computer interface for a tabletop system. It was shown that requirements like simultaneous interaction, multiple devices, resolution, workspace, ergonomics, and applications strongly influence the technical solution. It also became obvious that the research group ICVR employed different expertise in this research, not only for the system development, but also for designing adequate tasks and for performing user studies. The research introduced here was application-driven, focusing on an easy integration into daily business processes.

The examples also showed the necessity to decide, whether active or passive objects should be detected, meaning e.g. only TUI, or touch, or a combination of both. While active devices offer the possibility of transferring additional information, they might suffer from the fact that they are battery-driven. Passive devices on the other hand can be used all the time, but only have limited capabilities for transferring additional information. If this is required, the complexity of the underlying system (the table) would increase. It has to be noted that most passive systems do not distinguish between a passive TUI and a finger, which might lead to unwanted interactions with the system. Also new products like Microsoft Surface II (a Sony SUR 40 monitor with integrated IR-sensors) suffer from the fact that an ergonomic working is not possible, although the user can work in a sitting position. However, the screen simply detects touch and does not distinguish between fingertip, palm, or forearm, which could lead to misinterpretations by the underlying software. To avoid this, users behave unnatural when trying to hit the screen only in the point of interest. This is in particular uncomfortable for writing and sketching, for which we are used to rest hand and forearm on the table.

In summary, the ideal system is still not found yet, maybe because researchers look for an ideal solution that is capable of handling very different applications and requirements. However, this opens on the other hand new interesting research in this specific field of human-computer interaction.

REFERENCES

[1] Fjeld, M.; Bichsel, M.; Rauterberg, M.: "BUILD-IT: An Intuitive Design Tool Based on Direct Object Manipulation"; in: Wachsmut, I.; Froehlich, M. (Eds.): Gesture and Sign Language in Human-Computer Interaction; Lecture Notes in Artificial Intelligence, Vol. 1371; pp. 297 – 308; 1998

[2] Kunz, A.; Kennel, T.; Mazzone, A.; Müller, S.: "Workplace to Support the Teamwork in the Early Design Stages of the Product Development Process"; in: Proceedings of the Virtual Concept; Biarritz, France; 2002

[3] Ganser, C.; Kennel, T.; Kunz, A.; Birkeland, N.: "Computer Supported Environment for Creativity Processes in Globally Distributed Teams"; in: Proceedings of the ICED 2005; Sidney, Australia; pp. 109 – 110; 2005

[4] Ganser, C.; Steinemann, A.; Kunz, A.: "InfrActables: Multi-User Tracking System for Interactive Surfaces"; in: Proceedings of the IEEE Conference on Virtual Reality; Alexandria, Virginia, USA; pp. 253 – 256; 2005

[5] Hofer, R.; Nescher, T.; Kunz, A.: "QualiTrack: Highspeed TUI Tracking for Tabletop Applications"; in: Proceedings of Interact 2009; Uppsala/Sweden; pp. 332 – 335; 2009

[6] Hofer, R.; Kunz, A.; Kaplan, P.: "MightyTrace: Multiuser Tracking Technology on LC-Display"; in: Proceedings of the 26th SIGCHI Conference on Human Factors in Computing Systems"; Florence, Italy; pp. 215 – 218; 2008

[7] Hofer, R.; Naeff, D.; Kunz, A.: "FLATIR: FTIR Multi-touch Detection on a Discrete Distributed Sensor Array; in: Proceedings of TEI; Cambridge, UK; pp. 317 – 322; 2009

[8] Hofer, R.; Kunz, A.: "TNT: Touch 'n' Tangibles on LC-displays"; in: Proceedings of the 8th ICEC Conference; Paris, France; pp. 222 – 227; 2009

[9] Hofer, R.: "Tracking Technologies for Interactive Tabletop Surfaces"; Dissertation ETH Zurich, Switzerland; 2011

[10] Hofer, R.; Kunz, A.: "DigiSketch: Taming Anoto Technology on LC-displays"; in: Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems; Berlin, Germany; pp. 103 – 108; 2010

ISSN: 2236-3297