Touch & Write — A Multi-Touch Table with Pen-Input

Marcus Liwicki, Oleg Rostanin
DFKI
Germany
firstname.lastname@dfki.de

Saher Mohamed El-Neklawy
German University in Cairo
Egypt

Andreas Dengel
DFKI
Germany
andreas.dengel@dfki.de

ABSTRACT
In this paper we present a novel rear-projection tabletop called *Touch & Write*. It combines the FTIR technology for touching with the Anoto-technology for handwriting. This allows an implicit switch between the modes object manipulation, and content editing. Our system incorporates real-time gesture and handwriting recognition. Drawn objects and written concepts can be converted to digital information immediately. We introduce a functional application, the LeCoOnt concept mapping software makes use of the full capability of the *Touch & Write* table. Touching actions are used for arranging the concepts like sheets on a normal table, and to recognizes guestures like zooming. Pen-actions are used for drawing, connecting concepts, and handwriting. The handwritten strokes are automatically recognized and converted into a machine-readable string. This system provides a reliable alternative to common approaches which try to reconstruct the information from photographs.

Categories and Subject Descriptors
I.7.5 [**Document and Text Processing**]: Document Capture; B.m [**Hardware**]: Miscellaneous—*Design Management*

General Terms
Design, Algorithms

1. INTRODUCTION
In many discussion meetings the participants use paper to write down notes. Often, after some agreement, some of these notes are arranged to create a concept map. While this is a very fast way to generate information, there exists no fast, easy and reliable way to transform the data into a machine readable format and make it useful for other applications.

The common approach is to make a photograph of the discussion result, as illustrated in Fig. 1. On this picture document analysis methods are applied to recover the layout and recognizing the text. Another idea is to use cameras to record the handwriting already during generation in order to retrieve better recognition results [2]. While the discussion and brainstorming process in not disturbed by these methods, the result often is not very useful and much human processing is needed.

In this paper we propose an orthogonal approach to make the outcome of such discussion sessions available into an electronic format. Instead of improving the offline analysis and processing algorithms, we introduce a novel hardware which allows to add the content in a similar way as on paper, and additionally recognizes the data to make it available in the digital world. Figure 2 shows how the concept map of Fig. 1 is represented by out system.

1.1 Novel Input Metaphor
Almost three decades after the computer mouse started its triumph with the Apple Macintosh [1], it is about time for a next generation of Human-Computer-Interaction: Although robust and intuitive, the mouse does not completely reflect the interactions paradigms of the real world.

The challenge remained (almost) the same: Create input devices and software which is (even more) intuitive for first time users. This new generation seems to be found in Multi-Touch [6] tabletop environments, which provide hands on experience and offer a wider domain of usage scenarios. However, current Multi-Touch solutions lack in a way of intuitively switching between two important modes: moving the objects and editing their content, i.e., drawing or writing.

We have developed a novel rear-projection tabletop called *Touch & Write*. It combines infrared technology for the normal touching and moving with the digital pen technology for high resolution handwriting. This allows an intuitive switch between the modes object manipulation, and content editing. *Touch & Write* is an innovative new platform for creating applications, that users find natural to use. It seamlessly integrates the paper world into the digital world. Editing, arranging and writing tasks can be easily performed in an intuitive way.

The integration of touch and write functionality is motivated by creative or group tasks performed outside of the digital world. There, most people resort to using some form of pen and paper [8]. This is even the case when one tries to brainstorm or sketch out an idea, before putting it in

DAS '10, June 9-11, 2010, Boston, MA, USA

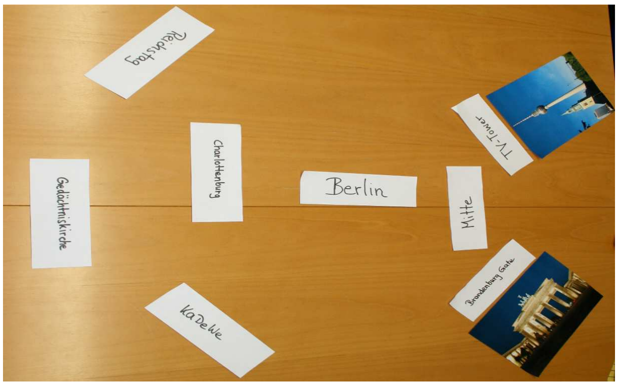

Figure 1: Concept mapping in the paper world

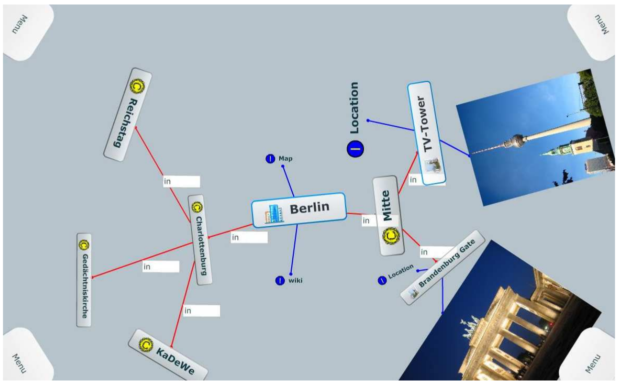

Figure 2: Concept mapping in the digital world

some digital form. Using pen and paper provides a hands on experience. Furthermore, it removes the limitation of thinking how to achieve a visual effect and gives a better chance to focus on the task itself. Such a concentration supports creativity when doing work, as there is more room for thought.

The rest of this paper is organized as follows. First, the technology is described with more detail in Section 2. Second, the implementation of the software services is briefly introduced in Section 3. Next, the novel concept mapping system is proposed in 4. Finally, a Conclusion is made in Section 5

2. TECHNOLOGY

2.1 Multi-Touch Environments

Touch environments are ones that allow users to directly interact with software using their fingers. Most commonly used of these provide only a single point of interaction, therefore having similar restrictions to that of a mouse. Examples of such could be found on notebooks' touch-pads, screen overlays and stylus driven tablet computers.

In recent years, newer technologies have appeared that enable multiple points of contact detection at the same time. These could be found in devices like the *Apple Trackpad*[1] or

[1]http://support.apple.com/kb/HT3448

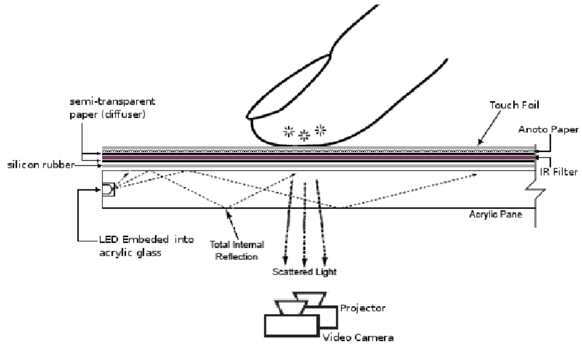

Figure 3: *Touch & Write* setup

the *iPod touch*[2], which allow for a limited number of simultaneous sensing points.

In contrast, other devices allow touch detection for arbitrary many points. Many of these have a camera based architecture. Here, the camera takes a live video capture of the touched surface. Computer Vision techniques are then applied on the video images (each frame), to obtain the points at which the user interacted with the surface (**blobs**).

2.2 Hardware Setup

Touch blobs are detected using FTIR, and the screen content is displayed via a rear projector. A description of the layers, shown in Figure 3, on top of the waveguide (acrylic glass) is as follows:

1. Scratch resistant touch foil is added to have a natural feel of the surface. This surface is further needed to reduce damage to the tabletop from the metal Anoto pen tip usage.
2. A sheet of Anoto dot pattern paper, to detect pen input.
3. A diffuser (semi-transparent paper) is added to prevent light from the projector to reflect back directly to the camera.
4. Infra-red blocking foil is used to further decrease background noise, usually occurring by sunlight.
5. Another diffuser.
6. A layer of *silicon rubber* is added on top of the waveguide. This is needed to simulate having oily fingers touching the surface. Such is required to get high quality blobs [3] seen by the camera. Furthermore, a small gap of air is needed between the silicon rubber and acrylic glass, to maintain total internal reflection when no touching of the surface is occurring.

2.3 Frustrated Total Internal Reflection

In order to provide an image for the camera that could be easily processed by software for blob extraction (see Section 3), we use the Frustrated Total Internal Reflection FTIR

[2]http://www.apple.com/ipodtouch/features/multitouch.html

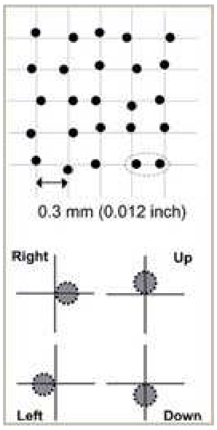

Figure 4: Dot Pattern

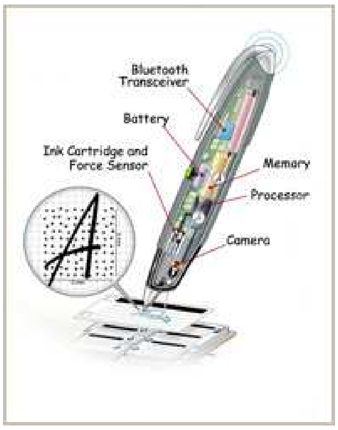

Figure 5: Digital Pen

technique [3]. This technique is usually used to identify the presence of touch on a surface. FTIR is based on total internal reflection, where light reflects within a medium (waveguide) when it is incident on its border at a critical angle. This phenomenon could be disturbed (**frustrated**) by having contact with the touch surface. Frustration causes light rays that were previously reflected inside the waveguide escape total internal reflection. Such a disturbance could occur at every point of contact. Hence, these light rays could be captured using a camera (see Figure 3) to know if a touch interaction has occurred or not. Furthermore, the amount/area of light is indicative of how much pressure has been applied to the touch surface. To distinguish light caused by a blob from background light, infra-red LED are used as the light that is reflected inside the waveguide. In addition, an infra-red filter is applied to the camera.

2.4 Anoto

Anoto[3] is a technology that facilitates the digitization of handwritten data. This is done by saving the stroke information. A strokes is mainly defined as a time ordered sequence of (x, y) coordinates, identifying the path of a pen while writing on a piece of paper. This information is gathered through a digital pen equipped with a camera that records (during writing) the Anoto dot pattern printed on the paper. The dot pattern and the digital pen could be seen in Figures 4 and 5.

Strokes data is transmitted live via bluetooth to other devices for processing. Such a possibility makes the Anoto technology well suited for use with a Multi-Touch environment. Furthermore, handwriting recognition methods could be applied to the stroke information.

3. IMPLEMENTATION

Our software builds on several standard technologies for Multi-Touch environments. We have enhanced some of the tools to deal with touch and write interactions simultaneously.

[3]http://www.anoto.com

3.1 TouchLib/Community Core Vision (CCV)

As discussed in Section 2, the image captured by the camera indicates where blobs are located. This needs to be processed for mapping the touch locations with the screen. After obtaining this data, it should be passed to the application which reacts to the touch interaction by the user.

TouchLib[4] and its newer version CCV are open source blob tracking applications that use Computer Vision techniques. This results in a collection of data regarding the blob such as:

- Location on the screen
- Area, which implies pressure
- Velocity of the blobs' motion

This information is packaged inside Tangible User Interfaces Object (TUIO) Protocol messages and sent to the application that is listening to those actions. TUIO [4] is a UDO-based communication protocol, used for data transfer to Multi-Touch applications. It provides a cross platform method for conveying the blob associated data.

The TUIO protocol facilitates a way of communicating blob data and further information concerning speed, pressure and duration. This makes it a very useful way for sending data that would be later processed by a gesture detection software. In addition, the protocol supports data about a blob's existence. From that information, it can be easily calculated how long contact was maintained with the surface.

3.2 Flex

All our written software is based on Adobe's Flash and Flex technology. The communication is done via ActionScript.

Flex is a Rich Internet applications development framework designed by Adobe Systems. It provides a layer of abstraction that is compatible with different browsers or platforms through the ActionScript Virtual Machine (AVM). Therefore, giving a chance for developers to create interactive application without having to worry about compatibility issues. The AVM is available for the internet in the form of flash player (could also be used offline), and for desktop environments through AIR.

When comparing Flash to Flex applications, it is hard to make the distinction purely from the output. This is specially because both compile to SWF binaries and run over the AVM. On the other hand, from a developer's point of view both are very different. The Flex framework offers many predefined components that make it easy to create an application. This is contrary to Flash, which is more animation and graphics oriented. Putting these points into consideration shows not only that Flash and Flex are not equivalent, but also that Flex is a better choice when it comes to creating applications. Flex applications could be written in two different tastes of syntax:

[4]http://nuigroup.com/touchlib/

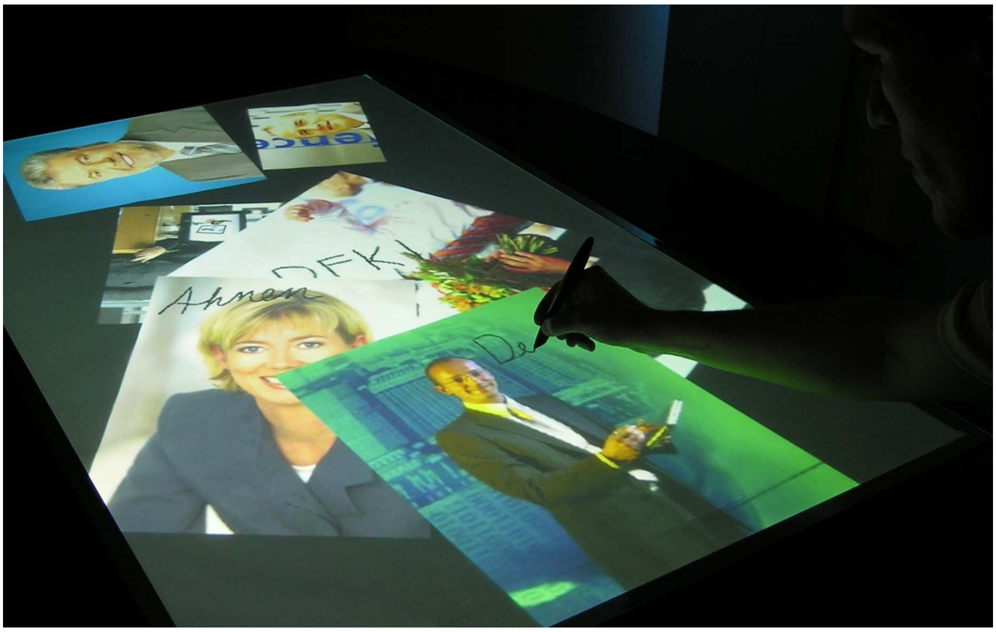

Figure 6: Image Annotator

- **ActionScript 3.0** is an object oriented language. It follows the ECMA Script scripting language standard. ActionScript also contains many event driven features. In addition, changes in a variable's state causes events to be fired, allowing different components to be easily tightly coupled. Such features make it possible to develop rich visual tools.
- **MXML** is a markup language mainly used to layout existing visual controls together. It is very similar to HTML.

Both ActionScript and MXML could be used interchangeably or mixed, by embedding ActionScript into MXML. To reach the final SWF binary, MXML is compiled to ActionScript classes, rendering them both equivalent.

4. CONCEPT MAPPING SYSTEM

We have implemented a concept map editor for Multi-Touch environments. Therefore, the existing concept mapping tool LeCoOnt[5] is enhanced from the traditional single user interaction to Multi-User Multi-Touch environments. Concept mapping sessions show both group and individual interaction possibilities. The former comes in the form of creating, editing, and viewing a concept map with the group. This gives room for discussion and collaboration between members, encouraging brainstorming of ideas and creativity.

Figure 1 shows how concept mapping is generally done in the paper-based world. Concepts are written down on sheets of paper. These sheets are arranged on a table according to the desired visual result. However, those settings usually suffer from transportation and re-use problems, i.e., it is not possible to re-use the outcomes from former sessions or to continue the brainstorming in another room.

Therefore, we have written a tool which transforms intuitive concept mapping into the digital world (see Figure 2). LeCoOnt is a concept map editing tool, which uses the full capability of the *Touch & Write*table. Touching actions are used for arranging the concepts and semantic zooming, while pen-actions are used for drawing, connecting concepts, and handwriting. The handwritten strokes are automatically recognized and converted into a machine-readable string by the MyScript Builder from Vision ObjectsVision Objects [5][6]. Having a computer application enabled for concept mapping gives the ability to digitally share acquired knowledge. Furthermore, since the task of creating a concept map is usually performed under a group setting, using *Touch & Write*as a platform fits perfectly.

In addition to the above mentioned functionalities it is also possible to view so called info items. Info items are attached to concepts for associating them with objects. Examples of such objects are documents, webpages, email address, images or geographical locations. Some applications provide the ability to view content that is not embedded in the current view, through opening it in another window. This could be in the form of another software that is able to perform this task. In many cases such an approach is unavoidable due to the limitation of the programming language or development framework used. On Multi-Touch environments, it is preferred to keep everything in the same viewing area without any window switching. The reason behind this is changing windows is usually done using keyboards and mouse, driving the interaction away from the natural one. Moreover, in a collaborative setting, change of view is distracting to other users, and would prevent them from a continuous work experience. The content of an attached info item could be viewed inside the application by enlarging it.

[5]http://lecoont.opendfki.de/

[6]http://www.visionobjects.com

Figure 7: Google Maps component

The information generated during a discussion session has a machine readable digital representation in the computer. We store not only the textual information but also the relations between the concepts and the concept maps.[7] This data can be imported to other application for re-use on the personal desktop [7].

4.1 Beyond Concept Mapping

We have implemented other software tools which which emphasize the importance of integrating the write technology. First, there is an Image Annotator which is used to view and organize photos (see Figure 6). This tool allows to write notes on the pictures using the digital pen. Finally the notes can be recognized and transformed into digital labels. One can easily imagine that not only pictures, but also documents which require some handwritten input can be edited in this application.

A second tool, namely the Google Maps component, allows users to view locations around the world. Navigation and zooming is realized using simple hand gestures. Creation of polygons and markers can be made with the digital pen and text can be added by writing it at the desired place. As a result, one has tagged geo-locations which were added as simply as marks on a paper-based map. An illustration of the application can be found in Figure 7.

5. CONCLUSION AND FUTURE WORK

In this paper we have presented a novel system for intuitively performing tasks with a pen on a screen. This *Touch & Write*System is a Multi-Touch rear-projection table which integrates touching actions and handwriting. This novel hardware gives previously unseen opportunities for a new generation of intuitive user interfaces. The construction of *Touch & Write* as a table top environment encourages collaborative settings.t' Users surround the table, discuss their ideas and work together. Since intuitive handwriting is allowed as an input metaphor, the creativity is furthermore supported.

Having this system at hand, we proposed a novel approach to make the outcome of brainstorming discussions and collaborative meetings available into an electronic format. This system incorporates handwriting gesture analysis and handwriting recognition to finally achieve an electronic representation of the generated information.

While the current system works quite well in praxis[8], we plan to investigate the performance and reliability of the system in the future. Furthermore, we will investigate other application areas, where online document analysis is expected to obtain better results than offline recognition. Currently we are designing a *Touch & Write*software which supports architects by recognizing and interpreting sketched construction plans and giving them useful hints and "lessons learned", where currently hours of searching in pictures and photographs are required.

Another research direction is to integrate new modalities like object detection and voice recognition into the hardware itself.

Acknowledgment

This work was supported by the "Stiftung Rheinland-Pfalz für Innovation" in Germany.

6. REFERENCES

[1] P. E. Bradley, M. J. Nuttall, and P. Alto. Computer mouse, 1987.

[2] G. A. Fink, M. Wienecke, and G. Sagerer. Video-based on-line handwriting recognition. In *Proc. 6th Int. Conf. on Document Analysis and Recognition*, pages 226–230, 2001.

[3] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 115–118, New York, NY, USA, 2005. ACM Press.

[4] M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza. TUIO - a protocol for table based tangible user interfaces. In *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW 2005)*, Vannes, France, 2005.

[5] S. Knerr, V. Anisimov, O. Baret, N. Gorsky, D. Price, and J. Simon. The A2iA INTERCHEQUE system: Courtesy amount and legal amount recognition for French checks. *Int. Journal of Pattern Recognition and Artificial Intelligence*, 11(4):505–548, 1997.

[6] NUI Group Authors. *Multi-Touch Technologies.* NUI Group, 2009.

[7] L. Sauermann, A. Bernardi, and A. Dengel. Overview and outlook on the semantic desktop. pages 1–19, 2005.

[8] A. J. Sellen and R. H. R. Harper. *The Myth of the Paperless Office.* MIT Press, 2001.

[7] the Resource Description Framework is used for storing the information http://www.w3.org/RDF/

[8] A demonstration video can be found on http://www.touchandwrite.de/