Authenticated Tangible Interaction using RFID and Depth-Sensing Cameras

Supporting Collaboration on Interactive Tabletops

Florian Klompmaker, Holger Fischer, Helge Jung
C-LAB
University of Paderborn
Paderborn, Germany
e-mail: florian.klompmaker@c-lab.de, holger.fischer@c-lab.de, helge.jung@c-lab.de

***Abstract*—Interactive large screen displays like tabletops or walls can enhance the interaction between humans and computers. A major topic is the collaboration between multiple simultaneous interacting people. However, most systems suffer from the problem that a distinction of different users is not possible. Hence, in this paper the authors present a work in progress approach of combining various existing technologies in order to enable personalized authenticated tangible interactions on a tabletop. Therefore, authentication using Radio-Frequency Identification in combination with depth-sensing cameras is used. We demonstrate the feasibility of the approach, the multiple advantages for interaction and give an outlook on further activities and lessons learned.**

Keywords-natural user interfaces; multitouch interaction; tangible interaction; interactive table; authentication

I. Introduction

Nowadays large multitouch displays like walls and tabletops can be found in a variety of use cases. Researchers have discussed the advantages like the direct interaction paradigm for a long time (e.g., in [1]). Because of their size, such displays are frequently related to multiuser and collaborative scenarios. It has been shown that large multitouch displays and especially tabletops are able to increase productivity and support the process of collaborative decision-making [2][3].

Additionally, in current research, many projects have identified the beneficial use of 'tangibles' (arbitrary physical objects which can act as input devices for various tasks). Tangibles that are placed on top of an interactive tabletop can be useful especially while performing complex tasks with even more complex tools, like a Computer-Aided Design (CAD) application [4]. Ishii and Ulmer have introduced tangible user interfaces to a broader community in 1997 [5]. They describe tangibles as physical objects that are graspable and act as computational input and output device at the same time. It has been shown that such kind of Human Computer Interaction (HCI) can improve working processes considerably. This is mainly achieved by giving physical form to digital information and by employing physical artifacts both as representation and as controls for computational media [1]. Associating input and output within the same interface improves the awareness and makes input devices more natural and intuitive since real world tools can be copied in shape and functionality.

However, multitouch displays and tabletops, that are able to detect and track fingers, as well as tangibles suffer from one disadvantage: They are not able to analyze 'who' is currently touching and interacting. However, user identification is desirable in collaborative working processes. It enables user authentication and, therefore, it is possible to apply different rights to different users. Further on, for researchers who are working on HCI analysis on tabletop devices, it is of particularly interest to track interactions along with information about the users that caused it.

Therefore, in this paper, we present a work in progress approach that uses Radio-Frequency Identification (RFID) technology and depth-sensing cameras in order to allow authenticated touch interactions and authenticated tangible interaction on a tabletop device.

We will first present related work before we introduce our approach. We will then discuss some lessons learned before we sum up.

II. Related Work

In this section, we present relevant related work that focuses on advantages of multitouch and tangible interaction for collaborative scenarios as well as on technical approaches for user authentication.

Considering topics like security and rights management personalized inputs and actions are desirable. Also, personal reasons are important since nobody likes to enter private data on a public visible virtual keyboard while anybody could be watching. Current systems are not able to distinguish between the authenticated user and the other users working on an interactive tabletop except for some technically complex installations.

The DiamondTouch approach from Mitsubishi Electric Research Laboratories is a multi-user tabletop device [6]. According to a collaborative work scenario the users can interact simultaneously without interfering each other. The surface, on which the image is projected, consists of a special layer with insulated antennas. The touches are transmitted via signals through the antennas. These signals are capacitivly coupled through the user and her/his chair to a receiver that identifies the areas of the table that are touched by her/him. A connected computer identifies the user and her/his touches and is able to interact based on this information.

 ISBN: 978-1-61208-177-9

The disadvantage of this approach is that only single touches are recognized by the system. In addition the users have to sit on a chair all the time. Hence, interacting around a tabletop is not really possible with this approach. The DiamondTouch can indeed distinguish between multiple users, but is does not enable a concrete identification or authentication. Further on tangibles cannot be detected by the device.

Some systems for people or item tracking are available on the market. However, the implementation of such systems involves extraordinary effort including setup and calibration. Regarding optical tracking some special patterns have to be attached to tracked objects (e.g., ARTrack [7]). Related to this, the users have to interact in a kind of artificial surrounding. Nor do such approaches allow authentication.

Regarding the tracking of tangibles, most tabletop devices use fiducial markers on the bottom of the object that are recognized by a camera. There are other technologies to detect objects in space (like the already mentioned ARTrack), not even just on a surface. Tracking technologies from the field of Virtual Reality for example are quite expensive, difficult to set up and mostly unmovable. There are approaches that introduce partially transparent patterns and therewith allow the stacking of tangibles [8][9]. Since stacking is recognized by the system more advanced interaction techniques on interactive tabletops are possible. Unfortunately here the use of opaque tangibles is not possible and the tangibles have to fit perfectly on top of each other requiring a very precise handling. The Gecko TUI uses magnets to create unique patterns [10]. These tangibles offer advanced interaction methods like detecting touch regions on the tangibles itself. It is shown that all these TUIs for advanced interaction require some kind of inner electronic circuit and/or mechanic making them less flexible and complicate to use.

Therefore we created the framework dSensingNI (**D**epth **Sensing** **N**atural **I**nteraction) [11]. It detects tangibles on a tabletop device and is able to analyze complex interactions like stacking, grabbing&moving and releasing. dSensingNI uses a depth-sensing camera and intelligent occlusion aware algorithms for image analysis. But, like all mentioned approaches, dSensingNI is not able to identify specific users and enable authentication.

Due to the missing authentication possibilities of existing systems, we present an approach that uses RFID technology as well as depth-sensing cameras to realize authenticated user interaction on an interactive tabletop.

RFID chips have been developed to track deliveries of goods. The chips can be either integrated in cards, stickers or directly in products. They are useful for the identification of items and also for the identification of their owner in case of an ID card. According to their production in bulk, RFID chips are a cheap possibility for a sophisticated kind of authentication, which is also conceivable for multitouch tables. Nevertheless RFID in combination with interactive tabletops are only mentioned in approaches concerning the identification of products and their presentation [12].

III. APPROACH

Our general idea is to combine OpenNI [13] based skeleton tracking using a depth-sensing camera with an RFID reader. After associating a read RFID tag to a user profile our vision algorithm tries to detect a skeleton's hand near the location of the RFID reader that is mounted beneath the table's surface (see Figure 2). Therefore, the algorithm is able to assign an RFID tag to a detected skeleton. The hand movements, touch events and tangible interactions of that user are 'authenticated' from then on. In the next subsection we will explain the mapping of touch events to users and afterwards the combination of skeleton tracking with dSensingNI for authenticated object interactions.

A. *Authenticated Touches*

A sketch of the view from the depth-sensing camera Microsoft Kinect on our local tabletop (called useTable [14]) is shown in Figure 1 (left). The task is to correct the perspective distortion caused by the camera not being located right above the center of the surface. The desired result is a mapping on the screen like shown in Figure 1 (right). Static configuration data provides the three-dimensional location of the four corners of the projection surface. From this information, a geometric plane E can be constructed by using one of the corners (e.g., TL) and vectors to the neighboring corners as already shown in Figure 1; d can be calculated by testing with known points of the plane (TL, TR, BL or BR).

$$E : TL + r \cdot \overrightarrow{g} + s \cdot \overrightarrow{h} = d \qquad (1)$$

The position of a hand as detected by the skeleton tracking framework is referred to point P – which is now being projected onto the plane E. To do this, the normal vector n = g × h of E is used to intersect a line from P into the plane along n:

$$\overrightarrow{n} \cdot P - d = P + t \cdot -\overrightarrow{n} \qquad (2)$$

By solving for t and calculating with the real values, we get PT – the projection of P into the plane E as if a shadow would be casted by the light coming from above the projection surface. By solving for r and s in equation 1 we get the logical position of projected point P on the plane - both margins need to be between 0 and 1 to be on the plane. By applying these factors to the real screen we can map the input from the skeleton tracking to a position on the actual projection.

The algorithm is very fast as the involved mathematics is not complex and the necessary computations can be done in only a few cpu instructions. The dominating factor regarding performance remains the tracking of the skeleton itself and, thus, the achievable framerate is only determined by the OpenNI tracking. In our tests, we could use the maximum framerate delivered by OpenNI in real-time without having to drop any frame. One optimization not even implemented by us is that some of the projection

 ISBN: 978-1-61208-177-9

calculations can be pre-computed before the capturing is started.

The algorithm presented above is applied on the incoming data to map hands that were detected by the skeleton tracking in the depth-image onto the actual presentation surface of the tabletop. This information is shown by the prototype in a separate window (see Figure 2 for an example screenshot).

While doing the visual tracking, the prototype uses an RFID reader to detect tags. For this task, the libtagReader embedded in the open-source project tageventor [15] is being used to periodically poll a list of present tags. This allows a fluid change of users since new ones can join the group at any time. When the visual tracking detects a hand in the area of the RFID reader, the software checks if there are tags present on the reader. All the tags are then assigned to the corresponding skeleton. The person is continuously tracked by the visual system until the person gets out of view. In that case she/he has to identify herself/himself again.

Currently, the prototype shows the assigned RFID tag identifiers alongside the tracked hands, but doesn't combine the multitouch input capability with the detected touch data from the useTable. We use the touch data from the table directly to track more precise touch events. Even though the depth-sensing camera can detect or estimate touch points on it's own [16]. Therefore, an interactive table is optional in this setup.

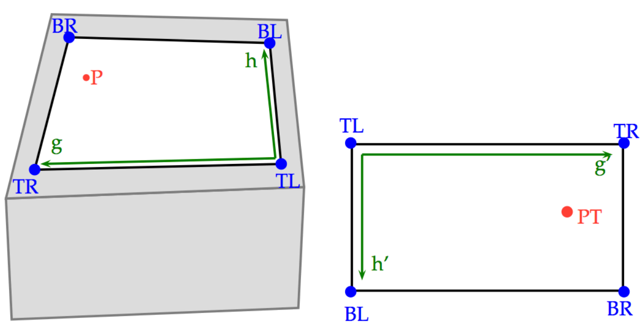

Figure 1: Symbolic representation of the projection surface as seen by the Kinect (left); Desirable target representation of a hand hovering over the projection surface (right)

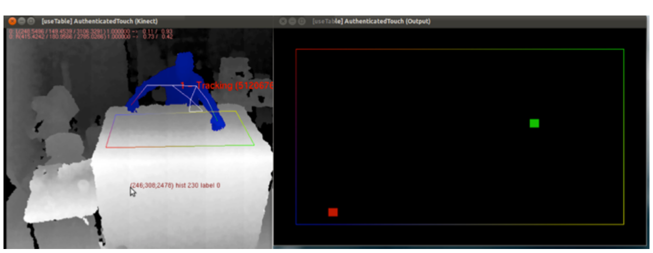

Figure 2: Authenticated touches on a tabletop device. Depth image with skeleton tracking activated (left) and detected finger touches that are assigned to two different users (right).

From the above, it is clear that – upon a multitouch event – the application can check if an authenticated user was detected in the relevant area of the projection surface and react upon this. A possible reaction could be to ignore an unauthenticated touch or allow special interaction only for selected persons.

B. Authenticated Objects

After realizing and testing the authenticated touch interaction we used the dSensingNI approach to detect objects on the useTable and their relation to each other [11]. dSensingNI is able to detect and track hands, objects and complex gestures. Objects are identified using their size and volume that can be analyzed from a depth image. Knowing which user is currently grabbing an object (using the information from the authenticated touch approach) allows the system to create user-object relationships and therefore record the user's interaction and assign personal objects to the users. In order to be able to detect even smaller objects we had to use a second depth-sensing camera (here again we used a Microsoft Kinect) as well as a second tracking computer. Since dSensingNI is able to transmit all detected object data using the TUIO protocol [17] the main application is easily extendable.

Since both our cameras for skeleton and for tangibles tracking use projected infrared patterns, they interfere each other in principle. Therefore we had to create a setup to avoid these interferences by using a top down setup for the tangibles tracking with dSensingNI and a front perspective for the skeleton tracking for user identification. Figure 3 (left) shows dSensingNI with a detected hand that just grabbed one out of three objects that are lying on the useTable. Figure 3 (right) shows a pointing gesture towards a tangible.

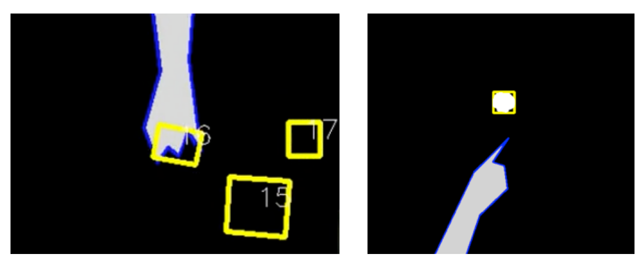

Figure 3: dSensingNI is used to detect objects, hands and grabbing gestures (left) as well as pointing gestures (right)

Besides detecting interactions with tangibles, dSensingNI is also able to detect freehand pointing gestures as shown in Figure 3 (right). This enables for example the selection of unreachable objects on an interactive tabletop. Using our combined approach this kind of selection is now also assigned to a specific user profile and therefore 'authenticated'.

IV. DISCUSSION

We implemented a simple map based application as a first prove of concept (see Figure 4). Here, we are able to apply authenticated touch so that only one predefined user is

 ISBN: 978-1-61208-177-9

allowed to manipulate the map (zooming and moving). Further on objects on the useTable are detected and assigned to the user who put it on the table or who performed the last interaction with the object.

In this setup, we had a few interference problems with the two depth-sensing cameras. In order to keep the skeletons in view we had to move the skeleton tracking camera a little towards the ceiling.

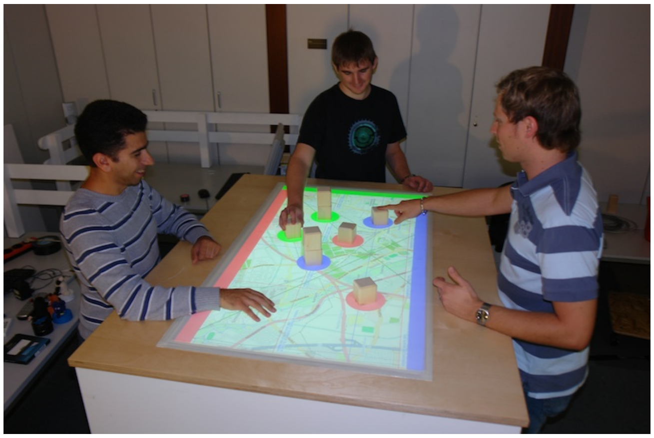

Figure 4: Authenticated collaborative tangible interaction on a tabletop

Compared with other solutions that recognize interaction in space, we presented a cost-effective approach that is also easy to set up. Even though our solution is not suitable for high-risk applications yet, because user IDs can be taken over by anyone in front of the depth-sensing camera quite easily, it demonstrates possibilities and scenarios for authenticated interactions on tabletops quite well.

Thinking of a tabletop system for collaborative mission planning in the field of disaster control management like that one we presented in [18], some workers might have additional rights in decision making others don't have. According to our approach users can authenticate themselves to the system and a specified set of interactions can be activated. Users can use tangibles or multitouch in compliance with her/his position in the organizational hierarchy.

V. CONCLUSION

In this paper, we presented a work in progress approach to enable personalized authenticated tangible interactions on an interactive tabletop device. Therefore, we used our dSensingNI framework that detects tangible objects on the surface of tabletops and evaluates complex interactions like stacking or grouping. In addition we extended this approach using RFID technology and a second depth-sensing camera for skeleton tracking. As a result, our prototype enables authenticated multitouch and tangible interaction for multiple users interacting on a tabletop. The interaction of multiple simultaneous interacting people can be traced and their steps in a solution finding process can be reviewed, e.g., in a HCI evaluation.

REFERENCES

[1] L. Yu, P. Svetachov, P. Isenberg, M.H. Everts, and T. Isenberg, "FI3D: direct-touch interaction for the exploration of 3D scientific visualization spaces". In: IEEE Transactions on Visualization and Computer Graphics, 16(6), 2010, pp. 1613-1622.

[2] E. Hornecker, P. Marshall, N.S. Dalton, and Z. Rogers, "Collaboration and interference: awareness with mice or touch input". In: Proc. Of the conference on Computer supported cooperative work, pp. 2008, pp. 167-176.

[3] V. Ha, K.M. Inkpen, T. Whalen, and R.L. Mandryk, "Direct Intentions: The Effects of Input Devices on Collaboration around a Tabletop Display". In: Proc. of IEEE International Workshop on Horizontal Interactive Human-Computer Systems, IEEE, 2008, pp. 177-184.

[4] R. Aish, "3D input for CAAD systems". In: Computer Aided Design, 11(2), 1979, pp. 66–70.

[5] H. Ishii and B. Ullmer, "Tangible Bits. Towards seamless interfaces between people, bits and atoms". In: Proc. of the Conference on Human Factors in Computing Systems, 1997, pp. 234–241.

[6] P. Dietz and D. Leigh, "DiamondTouch: A Multi-User Touch Technology". In: Proc. of UIST 2001, the 14th Annual ACM Symposium on User Interface Software and Technolgy, 2001, pp. 219-226.

[7] http://www.ar-tracking.de, 15.11.2011.

[8] T. Bartindale and C. Harrison, "Stacks on the surface: resolving physical order using fiducial markers with structured transparency". In: Proc. Of the International Conference on Interactive Tabletops and Surfaces, 2009, pp. 57-60.

[9] P. Baudisch, T. Becker, and F. Rudeck, "Lumino: tangible blocks for tabletop computers based on glass fiber bundles". In: Proc. Of the 28th international conference on Human factors in computing systems, 2010, pp. 1165-1174.

[10] J. Leitner and M. Haller, "Geckos: Combining Magnets and Pressure Images to Enable New Tangible-object Design and Interaction". In: Proc. of the Conference on Human Factors in Computing Systems, 2011, pp. 2985-2994.

[11] F. Klompmaker, K. Nebe, and A. Fast, "dSensingNI – A Framework for Advanced Tangible Interaction using a Depth Camera". In: Proc. of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (TEI), 2012, in press.

[12] Y. Huang, S. Wang, and Z. Deng, "An Interactive RFID-based Multi-touch Product Display System". In: Proc. of the International Multi-Conference on Complexity, Informatics and Cybernetics, 2010.

[13] http://www.openni.org, 15.11.2011.

[14] http://www.usetable.de, 15.11.2011.

[15] http://code.google.com/p/tageventor/, 15.11.2011.

[16] A.D. Wilson, „Using a depth camera as a touch sensor". In: Proc. Of the ACM International Conference on Interactive Tabletops and Surfaces, 2010, pp. 69—72.

[17] M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza, "TUIO - A Protocol for Table-Top Tangible User Interfaces". In: Proc. of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation, 2005.

[18] K. Nebe, F. Klompmaker, H. Jung, and H. Fischer, "Exploiting New Interaction Techniques for Disaster Control Management using Multitouch- , Tangible- and Pen-based Interactions". In: Proc. of the 14th International Conference on Human-Computer Interaction, LNCS 6776, Vol. 16, 2011.

 ISBN: 978-1-61208-177-9