Embodied Social Networking with Gesture-enabled Tangible Active Objects

Eckard Riedenklau and Dimitri Petker*

Abstract In this paper we present a novel approach for Tangible User Interfaces (TUIs) which incorporate small mobile platforms to actuate Tangible User Interface Objects (TUIOs). We propose an application of Tangible Active Objects (TAOs) in combination with gestural interaction for social networking. TUIOs represent messages while gestural input using these objects is used for triggering actions with these messages. We conducted a case study and present the results. We demonstrate interaction with a working social networking client.

1 Introduction

In the communication age digital exchange of information and keeping in touch with each other is getting more and more important. Social networks emerge for diverse kinds of communities and interest groups. Additionally gesture-enabled devices such as smart phones and smart pads allow the users to stay connected wherever they go. At work or at home, however, the connectedness may disturb the daily workflow. Embedding the interaction within the everyday environment may help making the user experience unobtrusive and ubiquitous [1]. In this paper we present a prototype of a system that embodies social networking in actuated gesture-enabled tangible objects.

Tangible Interaction is a subfield of Human-Computer Interaction (HCI). Researchers in this field search for new ways of interaction with digital information and functionality, keeping aloof from the traditional terminal consisting of display,

Eckard Riedenklau, Dimitri Petker, Thomas Hermann, Helge Ritter
Ambient Intelligence / Neuroinformatics Group, CITEC, Bielefeld University, Universitätsstraße 21-23, 33615 Bielefeld, Germany,
e-mail: [eriedenk,dpetker,thermann,helge]@techfak.uni-bielefeld.de

* This work has been supported by the Cluster of Excellence 277 Cognitive Interaction Technology funded in the framework of the German Excellence Initiative

keyboard, and mouse. This can be achieved by embodying those data in physical, graspable objects which users can interact naturally with using their everyday manipulation skills [2, 3].

Most TUIs use rigid, motionless objects, which only the user can manipulate. The system itself is unable to move those objects. Therefore researchers built actuated objects. Pangaro et al. [4] created the Actuated Workbench, a system which incorporates a grid of individually controllable electro magnets that enables ferromagnetic objects to be moved across a tabletop surface. Weiss et al. [5] elaborated on this technology to create a versatile set of widgets for interactive tabletops. Rosenfeld et al. [6] created actuated Tangible Objects differently, integrating small mobile robotic platforms into their objects to enable the system to save and restore arrangements of the objects on the interactive surface.

Gesture-based interaction is another hot topic in HCI research. It has frequently been applied to consumer products such as web-browsers or smart phones. Drawing shapes on a touch screen with fingers or the mouse triggers commands, such as ‘go back’ or ‘reload page’. As finding easily understandable gestures is not trivial, Wobbrock et al. already put alot of effort in collecting user defined gestures [7], evaluating different sets of gestures [8], and defined the guessability of such symbolic input [9]. RoboTable by Krzywinski et al. [10] enables the user to control mobile robots in a mixed-reality game scenario with motionless TUIOs. The authors claim that their system supports finger gestures or gestural input with passive TUIOs, but however unfortunately they do not explain if and how gestural input is used in their approach.

2 Tangible Active Objects and Tangible Desk

Our system is based on the Tangible Active Objects (TAOs) [11] which are used on the Tangible Desk (tDesk) (formerly known as Gesture Desk [12]). The tDesk is an interactive table equipped with a projector and a camera underneath a table-top glass surface equipped with projection foil, as shown in Fig. 1. We use it as a platform for interactive scenarios, such as multi-touch applications or TUIs. The TAOs are used as a TUI. They contain small low-cost robotic platforms which allow actuation of these tangible objects. The TAOs’ housing of these small robots are 3D printed cubes with an edge length of 5 cm ($\approx$ 2”). We attached visual markers underneath the TAOs for visual tracking, as depicted in Fig. 3. Like many small robotic mobile platforms, actuation is realized with a differential drive. An Arduino pro mini board[2] for rapid-prototyping of electronic systems controls this drive, whereas XBee modules[3] allow wireless communication and remote control. Because of the modular design of the TAOs they can be extended easily.

[2] `http://www.arduino.cc`

[3] `http://www.digi.com/products/wireless/point-multipoint/xbee-series1-module.jsp`

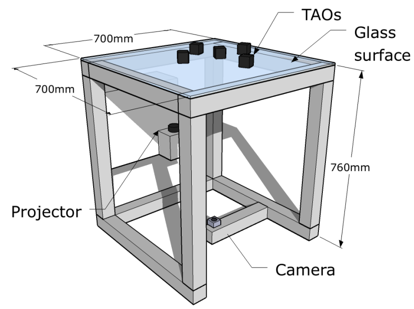


Fig. 1 The principal setup design. Our tDesk is equipped with a glass surface overlayed with a projection foil on which the projector mounted behind the table can project. A Firewire camera underneath the table allows visual tracking of the TAOs.

The software modules running on the host computer are organized in independent processes, communicating over the XML enabled Communication Framework (XCF) [13]. Fig. 2 depicts the software modules and their collaboration. A computer vision program analyzes the camera image from the Firewire camera mounted underneath the glass surface on which the TAOs interact to track markers derived from the reacTiVision markers [14] attached underneath the TAOs.

A path planning module takes the marker information and target requests of application modules, computes trajectories and navigation commands and navigates the TAOs. For this we adapted the potential fields apporach, described by Latombe [15]. Since the system has a complete overview of the scene, it is possible to compute attracting (targed position) and repelling (other TAOs) force fields for each TAO and navigate it via gradient descent through this force field. The corresponding navigation commands are relayed over the serial port to the wireless transmitter. To implement the embodied social networking application we added the following extensions to our system.

The system is intended to enable the user to physically interact with messages transmitted over a social network, such as Twitter[4] through gesture enabled TAOs. First of all, beside a speech synthesis module that can read interaction specific information to the user, we added back-projection capabilities to augment TAOs with visual information, such as messages, opened links, the different areas on the interactive surface and fields for textural input (via keyboard), etc. For this we replaced the previously used visual markers, to make the system more robust and independent from additional illumination that could be interfered by the projection. In a separate project we developed a self-luminescent visual marker based on infrared Light Emitting Diodes (LEDs). We arranged the LEDs on a 28×28 mm sized Printed Circuit Board (PCB) which exactly fits into the bottom of the TAOs body. In the top left corner of the PCB seven LEDs define the orientation of the TAO, other six LEDs encode its ID as depicted in Fig. 3(b)). To track these new markers we used the corner detection algorithm, proposed by Chen He [16]. After finding the corners of the markers it is possible to determine which of the ID encoding LEDs are illuminated, since their position relatively to the corner of the marker does not change.

[4] `http://www.twitter.com`

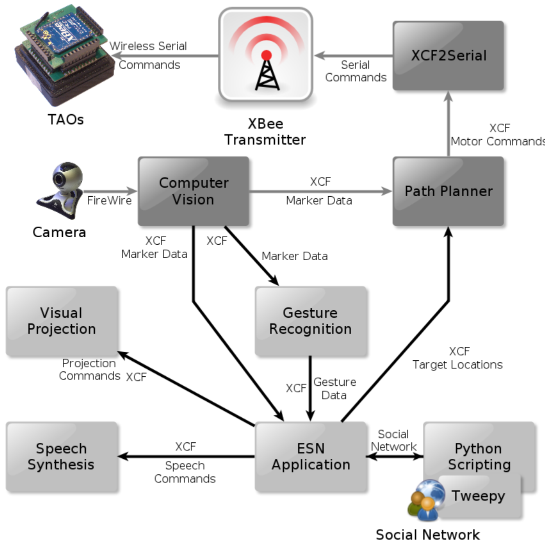


Fig. 2 This collaboration diagram describes the flow of information between the processes in our modular software architecture. The base modules used in this application are shown in dark gray, whereas the new modules are shown in light gray.

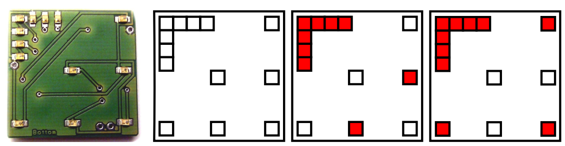


Fig. 3 Visual Marker: LED arrangement and example configurations.

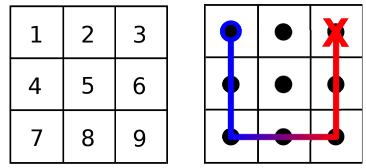
(a) The constructed PCB (b) Left: marker layout; middle and right: two of the 2^6 possible configurations

For triggering actions during interacting with social network messages we chose gestural input. Moving the TAOs along a specific path, thereby executing a 'gesture' as known from mouse gestures enables a novel and easy understood means to trigger actions. For initial experiments we utilize the gesture library LibStroke[5]. It provides very basic gesture recognition capabilities by segmenting gestures using a 3×3 grid of numbers as shown in Fig. 4(a). After performing the gesture shown in Fig. 4(b) LibStroke normalizes the tracking data to the grid and outputs "1478963". The output in terms of numeric sequences allows further processing of the detected gestures through simple string comparison. Because LibStroke is designed for a single mouse cursor as input, we extended the library to cope with multiple input devices such as the TAOs.

Interaction with those messages can then be implemented using TAOs as a handle according to the container concept introduced by Ulmer et al. in the mediaBlocks system [17]. Thereby every TAO can 'contain' such a message. To interface with a social network, we utilized the Python library Tweepy[6] which interfaces the Twit-

[5] `http://www.etla.net/libstroke`

[6] `https://github.com/joshthecoder/tweepy`



(a) Grid layout used by LibStroke (b) Example gesture

Fig. 4 Specification of Gestures in LibStroke; the example gesture (start at blue circle) results in the sequence "1478963"

ter API. Furthermore we added two new modalities. A display module enables the system to visually present content, such as the messages, opened links, the different regions of the interactive surface and fields for textural input (with a keyboard) etc. In addition, a speech synthesis module reads information to the user.

3 Case study

To investigate, which gestures users would expect to work with our embodied social networking client, we conducted an interactive case study. The subjects had to contemplate gestures according to a command selected randomly from a set of 11 commands. We asked the subjects to perform their gestures with one TAO, initially placed in the middle of the interaction area of the tDesk. In this study actuation is only used for automatically returning the TAO to the initial position after the subject finished the particular gesture to have the same initial situation for every trial. During the experiment we recorded the raw data of the trajectories and the output of the gesture recognizer. The gesture for each command was performed three times which results in a total of 33 trials per subject. During the experiment we also recorded the discussion between the subject and the experimenter for later analysis and transcription of gestures. After these trials we asked the subjects to fill out a small questionnaire to provide demographic information. Furthermore we asked our subjects if they already knew gesture-based interactions, e.g. mouse-gestures or finger-gestures and if they know and use social networks. We asked if the subjects could imagine to use such a system on their own desk and if they would accept standardized gestures or if they want an opportunity to define their own set of gestures.

4 Results

We conducted the study described in Section 3 with 15 subjects, all from Europe. All of them got instructions in their native language (German or English). 20% of the subjects were female. The average age of the subjects was 32.3, the youngest was 23, the oldest was 61 years old. All subjects were right-handed. 9 subjects already knew touch- or mouse-gestures but only 3 of them were actually using them. Social networks were known by 13 subjects and 11 of these subjects used them.

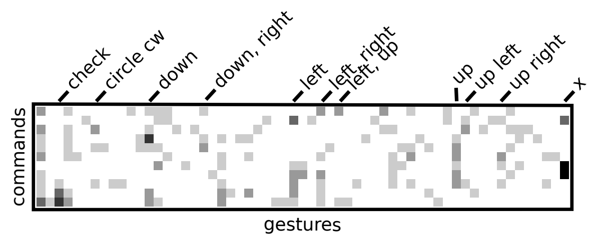


Fig. 5 All gesture occurrences plotted against the commands. The X axis represents all 60 different transcribed gestures, made by our subjects, the Y axis the corresponding commands (in the same order as in Fig. 6. The darker a pixel is, the more often the combination of command and gesture occurred. It is normalized for better visibility.

After the study we transcribed the gestures from the collected data (recorded trajectories, audio, and gesture recognition) since the gesture recognition did not work well for complex gestures. An overview over the complete data set collected in our experiments is visualized in Fig. 5. The plot is sparsely filled and there are few frequent command-gesture combinations visible as darker pixels. For better visibility we only consider the most frequently made gestures.

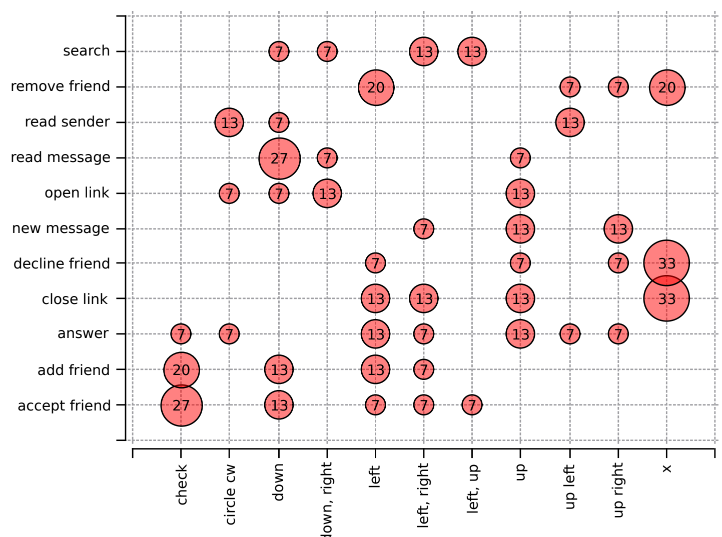


Fig. 6 This plot shows the (rounded) percentage of subjects that performed the gestures (*x*-axis) corresponding to the commands (*y*-axis). For a better overview we cropped away gestures with a score lower than 7%.

Fig. 6 depicts the most frequently occurring gestures performed by the subjects together with the percentage of subjects that chose the particular gesture in combination with the corresponding command. For the semantically similar commands such as *accept friend* and *add friend* the ‘check’ gesture was chosen most frequently, which is quite natural. Also the gesture ‘down’ occurs frequently. Subjects preferring this gesture described it pulling something to themselves. For the *answer* command the preferred gestures are ‘left’ and ‘up’, both are metaphorically meant as sending something back (opposite of reading direction or away from oneself). For the commands *close link* and *decline friend* the ‘x’ gesture was chosen most frequently, which is again quite natural. Also for the command *remove friend* this gesture was preferred beside ‘left’. For *new message* the most occurring gestures are ‘up’ and ‘up right’, where as the subjects came up with ‘up’ and ‘down, right’ most frequently for the command *open link*. The command *read message* got many correspondents in the gesture ‘down’. Subjects stated that this symbolizes the process of reading a text line by line. The preferred gesture for the command *read sender* was ‘circle cw’ which means ‘taking a closer look’. Some subjects stated that they located commands at corners or borders of the interactive surface and moved the TAO to one of these positions. This may result in different gestures if the starting point of the gesture is not located in the middle of the interaction area. One example is the *search* command. It found two winning gestures in ‘left, right’ and ‘left, up’. Here referencing with the border of the interaction area was crucial for the subjects. For a better understanding, the example trajectories of gestures from our collected data are depicted in Fig. 7.

Another interesting result of our study is that some subjects tried to make gestures in ways we did not think of beforehand. For example the TAO was turned in place, which was not recognized, because the gesture recognition only works on trajectories of 2D positions. Furthermore subjects tried to lift the TAO once or repeatedly as a metaphor for clicking or wanted to shake it. Obviously physical objects offer a much higher amount of flexibility for gestural commands.

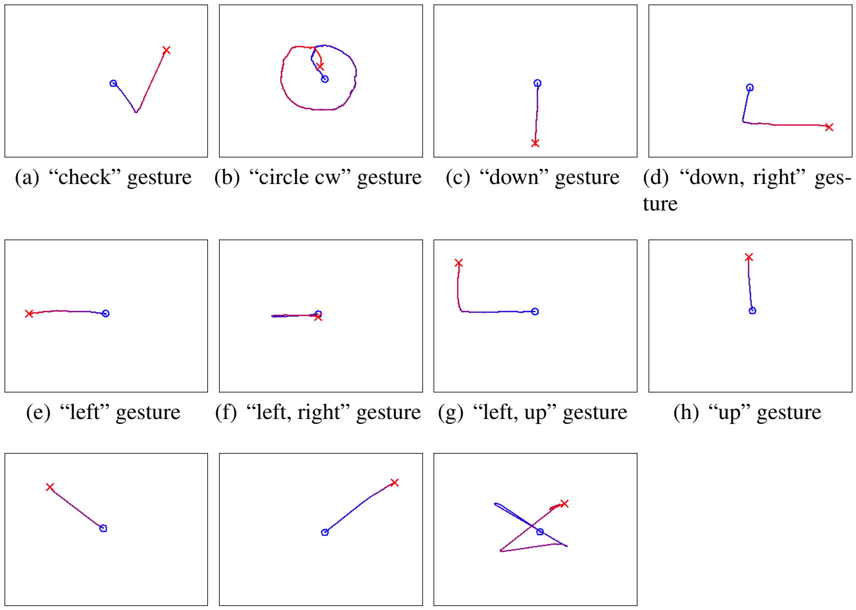
From the questionnaire data we found that 4 subjects stated that they would use it, 3 could not imagine using it and 8 were unsure. To the question if self-defined or standardized gestures were preferred, 9 subjects stated that they would prefer self-defined gestures, 4 would prefer standardized ones and 2 were unsure.

5 Interaction Design and Implementation of the interface

For our embodied social networking client we divided the table-top surface in four areas as depicted in Figure 8(a).

The actuation feature of the TAOs plays an important role in our application. Actuation is controlled by a finite state machine, which implements the state-graph depicted in Figure 9.

This results in the following behavior: Initially all TAOs are unassigned to any message and stay in the *waiting zone* until a new message is received from the social

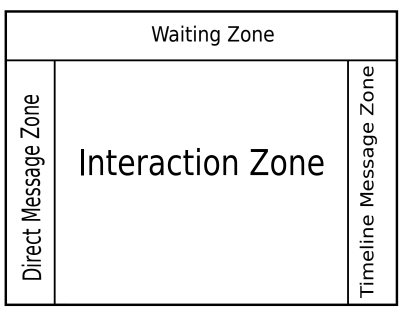


(a) “check” gesture (b) “circle cw” gesture (c) “down” gesture (d) “down, right” gesture

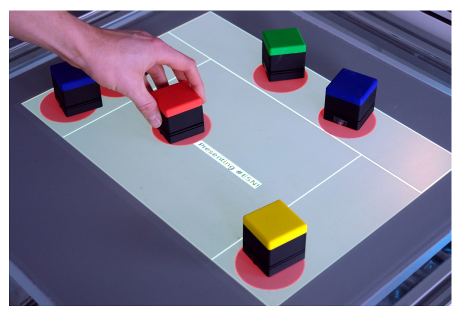
(e) “left” gesture (f) “left, right” gesture (g) “left, up” gesture (h) “up” gesture

(i) “up left” gesture (j) “up right” gesture (k) “X” gesture

Fig. 7 Visualizations of the winning gestures. A X symbol marks the start of the gesture movement, a circle symbol marks the end. The transition from start to end is represented by a color gradient from blue to red.



(a) Division of the tabletop surface: large interaction zone in the center, a message zone on each side and a waiting zone on top.



(b) Picture of the final system.

Fig. 8 Conceptional design of the interaction area and picture of running system.

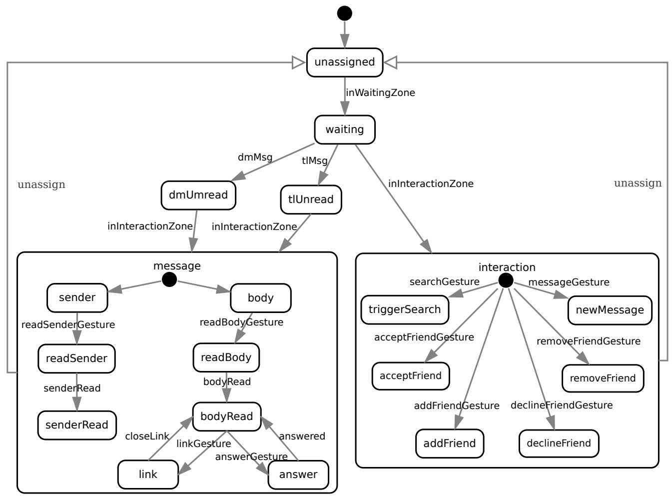


Fig. 9 State diagram: the implemented finite-state machine. Transition conditions contain message based events (such as *tlMsg*, indicating that a timeline message was assigned) and location based events (such as *inWaitingZone*, indicating that a TAO is in the waiting zone)

network. When a direct message (personal message from an other user) is received, it gets assigned to the leftmost TAO in the waiting zone. This TAO proceeds to the *direct message zone* area. If a timeline message (the timeline is a collected stream of postings from the user's friends) comes in it gets assigned to the rightmost TAO from the waiting zone. This TAO proceeds automatically to the *timeline message zone*. The TAOs in both message zones are ordered from bottom to top. The layout design of the interactive surface inherently maintains the chronological order of received messages as described later. If a TAO is taken out of the message zone and there are other TAOs above it, they automatically rearrange downwards. The user can take a TAO embodying a message from the two message zones and put it into the *interaction zone* to interact with the embodied message: through different gestures, the user can instruct the system to present the message (visually and through speech), to view the message's author profile, to open (or close) a link included in the message, or to open an input field for answering.

Additionally the user can put an unassigned TAO from the waiting zone into the interaction zone. In this case no message is assigned to the TAO so there are other interaction opportunities: opening an input mask for writing for a new message, adding or removing a buddy from the user's contact list or for searching the social network's history for a specific topic. An example picture of a user interacting with the system is shown in Fig. 8(b). A video demonstration of this application is provided on our website.[7]

[7] http://www.techfak.uni-bielefeld.de/ags/ami/publications/RPHR2011-ESN/

6 Conclusion

The presented approach makes a novel contribution to the HCI research field. As our first prototype of the user interface concept we built a social networking client, which combines actuated TUIOs with gestural input. We conducted a study to investigate which gestures are suitable for interacting with a social network with TAOs. We already found tendencies for suitable gestures and furthermore got valuable feedback from our subjects for improvements of our system design and considerations for our assumptions on the interaction design. To our experience gestural input is a useful way to interact with TUIs. However this needs to be empirically verified in user studies. We also learned that users would like to use the richer interaction possibilities that physical objects offer for performing gestures such as lifting, rotating or shaking a TAO.

For complex gestures, such as ‘x’ the gesture recognition was not robust enough with LibStroke so that we plan to utilize another custom recognition framework, such as the Ordered Means Models developed in our research group [18]. This will enable the user to use e.g. objects’ rotation in addition to the current translational gestures. For modalities that are not visually trackable, such as shaking (quickly) we need to create further extensions, such as additional integrated sensors. The modularity of our hardware and software makes such extensions easily applicable.

7 Acknowledgements

We thank Andreas Kipp, Matthias Schröder, and Dominik Weißgerber for helping creating the markers. We thank Christof Elbrechter for helping on the tracker and Dr. Dieta Kuchenbrandt for her advices on our study. We thank the German Research Foundation (DFG) and the Center of Excellence 277 Cognitive Interaction Technology (CITEC) who funded this work within the German Excellence Initiative.

References

1. M. Weiser, “The computer for the 21st century,” *Scientific American*, vol. 272, no. 3, pp. 78–89, 1995.
2. G. W. Fitzmaurice, H. Ishii, and W. A. S. Buxton, “Bricks: laying the foundations for graspable user interfaces,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co. New York, NY, USA, 1995, pp. 442–449.
3. H. Ishii and B. Ullmer, “Tangible bits: towards seamless interfaces between people, bits and atoms,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1997, p. 241.

4. G. Pangaro, D. Maynes-Aminzade, and H. Ishii, “The actuated workbench: computer-controlled actuation in tabletop tangible interfaces,” in *Proceedings of the 15th annual ACM symposium on User interface software and technology*, 2002, pp. 181–190.
5. M. Weiss, F. Schwarz, S. Jakubowski, and J. Borchers, “Madgets: actuating widgets on interactive tabletops,” in *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, ser. UIST ’10. New York, NY, USA: ACM, 2010, pp. 293–302.
6. D. Rosenfeld, M. Zawadzki, J. Sudol, and K. Perlin, “Physical Objects as Bidirectional User Interface Elements,” *IEEE Computer Graphics and Applications*, pp. 44–49, 2004.
7. J. Wobbrock, M. Morris, and A. Wilson, “User-defined gestures for surface computing,” in *Proceedings of the 27th international conference on Human factors in computing systems*. ACM, 2009, pp. 1083–1092.
8. M. Morris, J. Wobbrock, and A. Wilson, “Understanding users’ preferences for surface gestures,” in *Proceedings of Graphics Interface 2010 on Proceedings of Graphics Interface 2010*. Canadian Information Processing Society, 2010, pp. 261–268.
9. J. Wobbrock, H. Aung, B. Rothrock, and B. Myers, “Maximizing the guessability of symbolic input,” in *CHI’05 extended abstracts on Human factors in computing systems*. ACM, 2005, pp. 1869–1872.
10. A. Krzywinski, H. Mi, and W. Chen, “RoboTable: a tabletop framework for tangible interaction with robots in a mixed reality.” ACM, 2009, pp. 107–114. [Online]. Available: http://portal.acm.org/citation.cfm?id=1690388.1690407
11. E. Riedenklau, “TAOs - Tangible Active Objects for Table-top Interaction,” Master’s thesis, Faculty of Technology, Bielefeld University, Germany, June 2009, supervized by Dr. Thomas Hermann, Prof. Helge Ritter, and Tobias Großhauser.
12. T. Hermann, T. Henning, and H. Ritter, “Gesture desk - an integrated multi-modal gestural workplace for sonification,” in *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003 Genova, Italy, April 15-17, 2003, Selected Revised Papers*, ser. Lecture Notes in Computer Science, A. Camurri and G. Volpe, Eds., vol. 2915/2004, Gesture Workshop. Berlin, Heidelberg: Springer, 2004, pp. 369–379.
13. J. Fritsch and S. Wrede, *An Integration Framework for Developing Interactive Robots*, ser. Springer Tracts in Advanced Robotics, D. Brugali, Ed. Berlin: Springer, 2007, vol. 30.
14. R. Bencina, M. Kaltenbrunner, and S. Jorda, “Improved Topological Fiducial Tracking in the reacTIVision System,” in *Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*. Ieee, 2005. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1565409
15. J.-C. Latombe, *Robot motion planning*, 3rd ed., ser. The Kluwer international series in engineering and computer s ci. Boston [u.a.]: Kluwer Acad. Publ., 1993.
16. X. Chen He and N. H. C. Yung, “Corner detector based on global and local curvature properties,” *Optical Engineering*, vol. 47, no. 5, p. 057008, 2008. [Online]. Available: http://link.aip.org/link/OPEGAR/v47/i5/p057008/s1\&Agg=doi
17. B. Ullmer, H. Ishii, and D. Glas, “mediaBlocks: physical containers, transports, and controls for online media,” *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pp. 379–386, 1998.
18. U. Großekathöfer and T. Linger, “Neue Ansätze zum maschinellen Lernen von Alignments,” Master’s thesis, Faculty of Technology, Bielefeld University, Germany, September 2005, supervized by Prof. Helge Ritter and Dr. Peter Meinicke.