Towards Leaning Aware Interaction with Multitouch Tabletops

Khanh-Duy Le* [1], Mahsa Paknezhad* [3], Paweł W. Woźniak[1, 4], Maryam Azh[2], Gabrielė Kasparavičiūtė[1], Morten Fjeld[1], Shengdong Zhao[2], Michael S. Brown[3]

[1] t2i Lab, Chalmers University of Technology, Gothenburg, Sweden

[2] NUS-HCI Lab, National University of Singapore, Singapore

[3] School of Computing, National University of Singapore, Singapore

[4] VIS-MCI, University of Stuttgart, Stuttgart, Germany

khanh-duy.le@chalmers.se, paknezhad@comp.nus.edu.sg

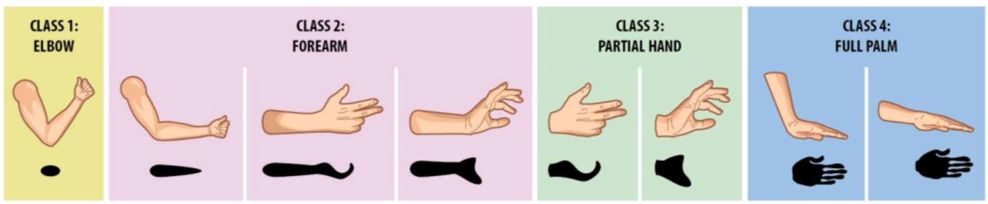

Figure 1: Tabletop lean posture classes (1-4): elbow, forearm, partial hand, and full palm (standing and seated). The black shapes are lean contacts between a user's body part(s) and a table's surface.

ABSTRACT

Interactive tabletops allow direct touch manipulation and recognizing simultaneous touch events. Users sometimes lean on the touch surface creating unintended touch input. Our work demonstrates how this unintended input can be employed to enhance interaction. In a study we develop a posture set organized into four classes. We present a vision-based machine-learning algorithm using an active shape model to recognize the classes. The algorithm categorizes lean gestures into one of the classes for interaction purposes. In a second study, we evaluate the model and propose interaction scenarios that use lean detection.

Author Keywords

Leaning; tabletop interaction; interactive surface; lean recognition; active shape models.

ACM Classification Keywords

H.5 Information interfaces and presentation (e.g., HCI) (I.7), H.5.2 User interfaces (D.2.2, H.1.2, I.3.6).

INTRODUCTION

Interactive tabletop surfaces allow direct manipulation of the interface by recognizing touch events. While direct-touch input typically happens at single or multi-finger level [11], users tend to lean on the table using different parts of their arm while interacting with multi-touch surfaces [2, 9]. Such leaning behavior often occurs as a result of distal object selection, fatigue [20], or the need for extra support during high-precision tasks [16]. However, this behavior is often not processed very well by current interfaces, thus reducing accuracy and usability [6].

A common approach to solve this problem is to detect and discard multitouch contact points above a set threshold [2, 18]. Besides simply offering binary lean detection information, such approaches discard all information from unintended contact postures that could, potentially, be used to improve user interaction with tabletops. In particular, we believe that lean contact information can be employed by the interface for adaptive decision making because it contains information about the presence, posture, and even emotional state of a user.

In this paper we first present a user study conducted to understand different leaning postures (referred to as *lean postures*) users take while performing common tasks on tabletops, in order to generate a *posture set* for unintended lean contacts with the interactive surface. We then present an Active Shape Model (ASM)-based [1] algorithm to recognize four classes of lean postures. The algorithm also yields information on the user's position in relation to the surface and handedness of the posture. Finally, in a second study, we evaluate the model and propose possible scenarios that illustrate its utility.

NordiCHI '16, October 23-27, 2016, Gothenburg, Sweden
© 2016 ACM. ISBN 978-1-4503-4763-1/16/10…$15.00
DOI: http://dx.doi.org/10.1145/2971485.2971553

* Khanh-Duy Le and Mahsa Paknezhad are joint first authors and contributed equally to this work

RELATED WORK

To process unintended contact as undesirable touch events, designers are advised to consider gestures that are suitable for interaction with multitouch devices, and to avoid gestures that require unnecessary arm movements and tapping actions [20]. These include asking the user to identify the tabletop workspace before interacting with the device [6, 13], or considering only oblique touches as user input [8]. Other techniques reject them [15] or adjust the interface based on the unintended contact areas [3, 7, 16].

Other approaches leverage contact areas larger than the fingertips as alternative input techniques. For instance, contact areas of different parts of the hand were proposed as new input streams [17], or as means of expanding the range of expressive input gestures on multitouch surfaces [7, 19, 20]. Empirical results show the potential utility of those alternative gestures in tabletop interaction scenarios [5, 10].

Hand posture sets for multitouch interactions are only conceptually proposed [4, 17]; and empirical usability studies rely on Wizard of Oz techniques [5] to simulate posture recognition, or utilize thresholds to distinguish between small and large contact areas [16]. Other tabletop gesture recognition systems depend on external hardware and more complex setups [8, 18, 19]. Existing built-in depth sensor images are often used for detecting multitouch rather than hand gestures [15]. To our knowledge, there has not been any work attempting to recognize and categorize lean contacts to enhance user interaction on tabletops.

STUDY 1: DEVELOPING A LEAN POSTURE SET

We conducted an explorative user study to understand common lean postures that occur while interacting with tabletops. We used the data obtained to build preliminary insights on how lean postures differ.

We simulated a tabletop display using a glass panel covered with thin paper through which light could pass. The contact areas could thus be recorded using a depth camera located beneath the table. The study involved four users (three female; two left-handed; mean age=28). User behavior was video recorded while performing two tasks (solving a tangram puzzle and sketching), under two conditions (standing and seated). The users performed three trials per task. The sketching task was performed with water-based artistic materials to simulate one finger input, while the puzzle task required single or bi-manual interactions on small objects. The order of tasks was counterbalanced among participants.

Two researchers worked together to manually analyze and classify all the captured touch data on the table surface based on their size, length and shape differences as well as the body parts generating them, and related the data to corresponding behaviors in videos to determine a set of lean posture categories. The result was then discussed and agreed with other two researchers. Our analysis yielded four distinct classes based on the significant differences in their shapes and the body parts generating them: elbow (Class 1), forearm (Class 2), partial hand (Class 3), and full palm (Class 4) as shown in Figure 1. There are internal shape variances within Classes 2, 3 and 4. Within Class 2 shapes vary with how much a user's forearm contacts the table's surface. Within Class 3, shapes depend on how a user's hand rests on the table. Within Class 4, standing gives a slightly smaller shape area than sitting. It should be noted that we observed lean postures involving varying number of fingers, but chose not to include these in this work due to high complexity.

PROOF-OF-CONCEPT ALGORITHM

In this section we propose a proof-of-concept vision-based algorithm as a first attempt to recognize lean postures. We employed ASM for our model since this is established practice in the computer-vision community. We developed four ASMs with respect to four lean posture classes; each going through the following steps (see Figure 2):

- Thirty sample images of each lean posture category per class were selected from the training dataset after flipping the depth image to represent the same hand (right hand).
- The samples were placed at the center of a global coordinate system and were rotated to have a similar vertical angle. Seventy contour points were manually defined around the contact area of each sample.
- An ASM was trained for each class from the samples of the categories they cover.

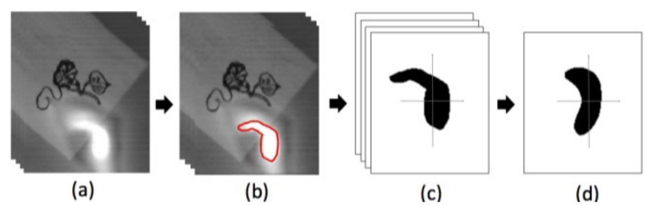

Figure 2: Modeling steps for each lean posture: (a) sample images of the lean posture, (b) defined contour points on vertically aligned images, (c) binary representation of the lean contact area, and (d) the mean model for the lean posture.

The four ASMs were applied to the lean contacts in 10 iterations to classify observed lean contacts. We then compared the alignment results of each of the four models to the contact area using normalized cross correlation (NCC) [12]. The model and its corresponding class with the best NCC score exceeding a given threshold (0.3) was selected.

A user's handedness or position around the tabletop could be inferred from the lean posture. The potential table side (sides) of the user was determined by measuring the greatest planar distance from the touch area to each side of the tabletop and discarding the sides further than the length of the whole forearm, taken from the training dataset. For Class 2, the left- and right-handedness of the touch is determined by the angle of the touch area with regards to the tabletop's coordinate system and the potential location/side of the user.

For Classes 1, 3, and 4, the interaction area of the user needs to be considered to determine handedness of the contact. The reason is that the *elbow* with its round shape only conveys limited rotational information unless a small part of the forearm is also in contact with the screen. The angle of the detected posture in the *partial hand* class and *full palm* class did not perform well for determining handedness, due to the shape of the model.

STUDY 2: EVALUATING THE MODEL

We built an application running on a horizontally placed Samsung SUR40 (40-inch screen size), which uses PixelSense (infrared-based) technology to capture images of users' touch interactions on the screen. Sensors in the individual pixels in the display register what is touching the screen at two frames per second (FPS). Our informal observation in study 1 revealed that users did not change their lean contacts too quickly or too often in a short period of time. Hence, this capture frame rate is sufficient to acquire necessary and skip unnecessary data. The application was remotely controlled by an Android application to avoid distracting participants during the experiment. An RGB camera placed 50 cm in front of the table also recorded user behavior on and above the table.

To collect data to train and evaluate the model, we recruited 20 participants (5 females), averaging 24.5 years old (SD=3.02) in the university campus and nearby companies. 43% of the participants had prior interaction experience with interactive tabletops. Users were seated to perform three tasks (typing, sketching and tangram puzzle solving) on a table. They performed a typing and sketching task individually and were paired to solve the tangram puzzle. These tasks were chosen considering current tabletop applications (maps, photos, browsing, and drawing), natural posture of a user while working at table (leaning) and vision of tabletop applications in daily use in the future (office works like writing/drawing, typing or multi-user collaboration).

All the artifacts presented to participants within the experiment such as drawing paper, simulated keyboard (keyboard image printed on a paper) and tangram puzzle pieces were made from transparent plastic foils (Figure 3a,b,c). This helped avoid the occlusion of physical artifacts on the touch interaction of users in the captured infrared images (Figure 3d). Captured depth images in this study were used to train and test the algorithm using a leave-one-out strategy.

The algorithm was tested on a 64-bit operating system PC with an Intel® Core™ i7-2600 CPU and 16GB of RAM. Table 1 shows the algorithm's accuracy and mean recognition time. The best accuracy at 94% was achieved with Class 4. However this lean class often appeared when users were idle, hence not much meaningful information to be utilized as an input. The relative lower accuracy observed for class 2 was due to those leans with a smaller contact area compared with a longer shape used in the training phase. Class 1 and Class 3 had a similar acceptable recognition rate. Our observation also revealed that Class 1 and Class 2 appeared more frequently than the others. There are two cases where the algorithm could not detect any lean although leans did appear in the image contents. Conversely, the algorithm wrongly recognized leans in five test images whose content did not contain any lean. The algorithm performed similarly for about one second for all classes. We also observed that lean contacts did not change quickly while a user was performing a task. For example, when sketching and typing, users often rested their forearm or part of the hand stably and just mainly used their fingers to perform the task (typing or controlling a pen to draw). Similarly, in the tangram puzzle, lean contacts changed at a rather low frequency; we assume to avoid fatigue. Additionally, any decision made in response to lean contacts is not critical to the user's task at hand. Since the posture is unintended, the recognition method also only facilitates task performance in non-obtrusive ways. Thus, intentional system or interface adaptiveness in (near) real-time is not necessarily the goal of the proposed method. A recognition time of about one second seems acceptable for the scenarios we studied here.

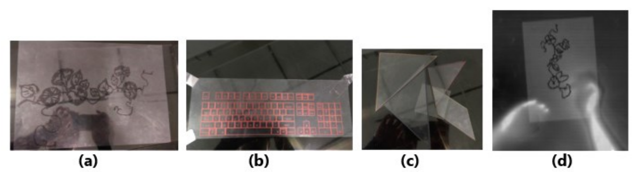

Figure 3: Tabletop workspace examples with (a) transparent foil in drawing task, (b) keyboard printed on transparent foil, (c) transparent puzzle pieces, and (d) infrared image captured by tabletop pixel-level sensor in the experiment.

DISCUSSION AND OUTLOOK

Recorded data (see Figure 4) helped us observe cases where users implicitly and unintentionally formed their personal territories by different combinations of lean postures. In future work, these combinations of postures and corresponding personal spaces could be recognized based on lean contact information such as location, orientation, category, handedness, boundaries, etc. which were shown to be extractable by our proof-of-concept, computer-vision based algorithm.

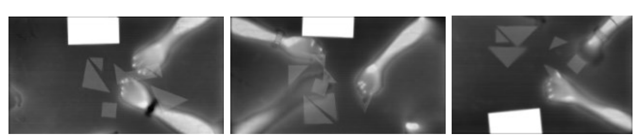

Figure 4: Tabletop workspace with (left) two forearm leans framing a closed personal space, (center) right user creating an open space with forearm and elbow, and (right) two forearm lean contacts forming a partially open personal space. Bright regions are paper sheets presenting the task; such shapes are not yet recognized by our method.

We also expect that leveraging lean postures may contribute to tabletop territory management [14]. In particular, Class 1 and Class 3 may help in dynamically determining personal territories on the table in unobtrusive ways rather than asking a user to explicitly do that [13]. Hence, subtle cues can be offered to users such as intelligent grouping of personal and shared items, adjusting content to make it more visible to the users, acknowledging user's presence, or adaptively showing and adjusting tools/widgets suitable to the current task. Moreover, information extractable from lean contacts might also help recognize the current posture of the users, which can be combined with conversational gesture theories to better understand and support users and their relation in a collaboration session such as making them more active or engaged in the collaboration.

		Recognition output						
		Class 1	**Class 2**	**Class 3**	**Class 4**	**No lean**	**Accuracy**	**Recognition time (s)**
Image content	**Class 1**	34	0	6	0	0	85%	0.99±0.02
	Class 2	16	102	11	12	1	72%	0.98±0.04
	Class 3	5	1	46	2	1	84%	0.99±0.02
	Class 4	1	1	0	32	0	94%	0.99±0.03
	No lean	0	1	3	1			

Table 1: Performance results of the recognition algorithm for each lean class tested with data from study 2.

ACKNOWLEDGEMENTS

Thanks to STINT (2012-019), CELTIC (C2013/1-3), DFG/SimTech Excellence Cluster (EXC 310/2), Adlerbertska, Philippa Beckman, and Ayça Ünlüer.

REFERENCES

1. Ahmad, T., Taylor, C.J., Lanitis, A. & Cootes, T.F. Tracking and recognising hand gestures, using statistical shape models. *Image and Vision Computing*. (1997), 15(5), 345–352.
2. Annett, M., Gupta, A. & Bischof, W.F. 2014. Exploring and Understanding Unintended Touch during Direct Pen Interaction. *TOCHI '14 (2014)*. 21(5), 1–39.
3. Brandl, P., Leitner, J., Seifried, T., Haller, M., Doray, B. & To, P. Occlusion-aware menu design for digital tabletops. In *Proc. CHI '09 EA* (2009), 3223–3228.
4. Epps, J., Lichman, S. & Wu, M. A study of hand shape use in tabletop gesture interaction. In *Proc. CHI '06* (2006), 748–753.
5. Freeman, D., Benko, H., Morris, M.R. & Wigdor, D. ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proc. ITS '09* (2009), 165–172.
6. Gerken, J., Jetter, H.-C., Schmidt, T. & Reiterer, H. Can "touch" get annoying? In *Proc. ITS '10* (2010), 257–258.
7. Leithinger, D. & Haller, M. Improving Menu Interaction for Cluttered Tabletop Setups with User-Drawn Path Menus. In *Proc. TABLETOP '07* (2007), 121–128.
8. Marquardt, N., Kiemer, J., Ledo, D., Boring, S. & Greenberg, S. Designing user-, hand-, and handpart-aware tabletop interactions with the TouchID toolkit. In *Proc. ITS '11* (2011), 21–30.
9. Morris, M.R., Everitt, K. & Ryall, K. and Forlines, C., Chia Shen Experiences with and Observations of Direct-Touch Tabletops. In *Proc. TABLETOP '06* (2006), 89–96.
10. Nacenta, M.A., Kamber, Y., Qiang, Y. & Kristensson, P.O. Memorability of pre-designed and user-defined gesture sets. In *Proc. CHI '13* (2013), 1099–1108.
11. O'Hara, K. Interactivity and non-interactivity on tabletops. In *Proc. CHI '10* (2010), 2611–2614.
12. Sarvaiya, J.N., Patnaik, S. & Bombaywala, S. mage Registration by Template Matching Using Normalized Cross-Correlation. In *Proc. ACT '09* (2009), 819–822.
13. Schmidt, D., Chong, M.K. & Gellersen, H. IdLenses: dynamic personal areas on shared surfaces. In *Proc. ITS '10* (2010), 131–134.
14. Scott, S.D., Sheelagh, M., Carpendale, T. & Inkpen, K.M. Territoriality in collaborative tabletop workspaces. In *Proc. CSCW '04* (2004), 294–303.
15. Vogel, D. & Casiez, G. Hand occlusion on a multi-touch tabletop. In *Proc. CHI '12* (2012), 2307–2316.
16. Wigdor, D., Benko, H., Pella, J., Lombardo, J. & Williams, S. Rock & rails: extending multi-touch interactions with shape gestures to enable precise spatial manipulations. In *Proc. CHI '11* (2011), 1581–1590.
17. Wilson, A.D., Balakrishnan, R., Hinckley, K., Hudson, S.E. & Xiang Cao ShapeTouch: Leveraging contact shape on interactive surfaces. In *Proc. TABLETOP '08* (Oct. 2008), 129–136.
18. Wobbrock, J.O., Morris, M.R. & Wilson, A.D. User-defined gestures for surface computing. In *Proc. CHI '09* (2009), 1083–1092.
19. Wu, M. & Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proc. UIST '03* (2003), 193–202.
20. Yee, W. Potential Limitations of Multi-touch Gesture Vocabulary: Differentiation, Adoption, Fatigue. *Human-Computer Interaction. Novel Interaction Methods and Techniques*. (2009), 291–300.