Foot Position as Indicator of Spatial Interest at Public Displays

Bernd Huber
Korea Institute of Science and Technology
Seoul, South Korea
berndbhuber@gmail.com

Copyright is held by the author/owner(s).
CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
Motivated by and grounded in observations of foot patterns in a human-human dialogue, this study explores expressions of spatial interest through feet at public displays. We conducted an observation and recorded user foot orientation and position in a public information display environment leading to data about 84 interaction sessions. Our observations show that characteristic foot patterns can be matched with two user intentions: (A) Users who seek access to specific information, and (B) users who don't seek specific information. With the goal to detect intention through foot patterns, we classified characteristic foot patterns with a *SVM* pattern recognition algorithm, which resulted in a detection accuracy of 84.4%. This work can be valuable for researchers designing context-aware public displays.

Author Keywords
Public displays; Unconscious interaction; Embodied interaction; Context of use.

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

Introduction

Public displays need to identify the context to meet user needs. Many studies in this domain are based on the proxemics theory of interpersonal spatial relationships introduced by Edward Hall [1].

Public display interaction frameworks can be divided into ad-hoc and observation based models [4]: As an example of an ad-hoc model, the interactive public ambient display [5] interpreted proximity and body orientation as a direct signal for display content adaptation. Similarly, the proximity toolkit [6] used proximity to control a multimedia-application, with a focus on the head orientation. In observation-based models, such as the *CityWall* [7] and the *Opinionizer* [8], the observation focus is the social behavior of the audience. We show that measuring foot pattern is a concrete approach for user intention recognition.

To measure user behavior concretely, a variety of sensors have been developed. Eye gaze was studied as a measure of attention since it is directly related to visual attention. The gaze tracker in [8] is a vision-based gaze tracking system that captured the visual attention of users through eye patterns. The vending machine in [9] could estimate its users' age and gender through a vision tracking system and gave drink recommendations accordingly. Similar to public displays, *Medusa* located users around a tabletop computer and associated touches with them using a multitude of proximity sensors [10]. Similar to our laser scanner approach, *Multitoe* enabled high precision tracking of feet using a back-projected floor [11]. This approach, however, did not consider user intention.

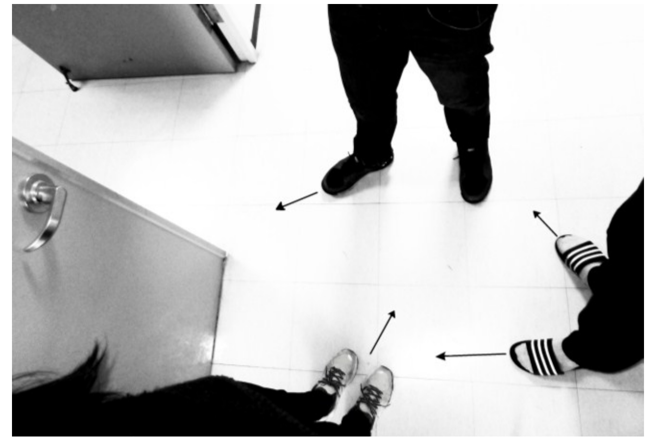

Figure 1. Foot patterns as a measure for user intention. The direction of the feet is related to spatial interest.

In this work, we observe the user's foot position, a type of body language which expresses information about user intention. More specifically, feet are used for walking and stable platforms and so can be used for pointing, as with other parts of the body [2, 3]. Therefore, feet can be used as a subconscious indicator of spatial interest and they may send a signal about the things in which people are interested in at public displays. With this motivation, we present a new approach: Our contribution is a foot pattern measurement used for intention detection.

We define foot pattern features using the angle of each foot and their relative position. Because standing feet are located in one plane, the features can be extracted from one-dimensional depth data. In our observation, we use a laser scanner which can be used under sunlight, to capture foot patterns. The implementation of this simple sensing system for foot patterns could

Figure 2. The default screen is the local map, which is the display's main purpose.

enhance public display interaction by addressing the audience according to the user context.

Observational Study

Observation Setup

To observe the behavior of a public audience, it was necessary to test the application of a display in a real world environment [12]. We selected the entrance lobby of a research institute as the testing environment with about 100 passersby per day. To address the audience's needs, we designed an information system for the research institute with map, information videos and event schedule as well as general information such as weather and time. We designed a common graphic user interface (GUI) in the public display system with menu buttons at the bottom and the content displayed in the center (Figure 2). We selected a 40" multi-touch display with a PQ-Labs IR touch frame attached. We placed the display next to a magazine holder and a water fountain to increase the interest of those passing by [12]. Each static foot pattern was stored in a log file. The foot pattern data combined with video recordings and touch log formed the foundation of our observation results. Applying a system that combines entertainment (draw bystanders with video function) and urgent needs (map) was beneficial for intention differentiation during the observation.

Device Setup

To record foot pattern data, we installed a Hokuyo URG laser rangefinder at the bottom of the display; the laser rangefinder scanned 2 cm above the floor (See Figure 3). The scanner captures objects at distances from 6cm to 4m at a rate of 30 Hz. We tracked feet using background subtraction [13]. Our algorithm determined three feature points for each foot: the forward section of the foot and its two edges (see Figure 4). Only the nearest feet were tracked.

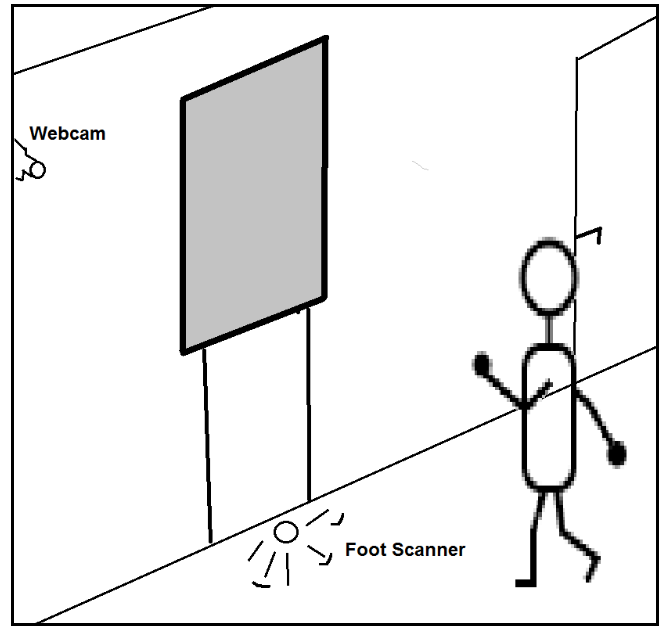

Figure 3. The laser rangefinder is positioned at the bottom of the display. The maximum range of the scanner was limited to 1.5 m.

Using the coordinates of each point, we defined the foot angle as the angle of the longest edge relative to the horizontal display alignment. A brief technical evaluation showed an angle accuracy of ±10° and a position accuracy of ±5 cm (of the foot peak).

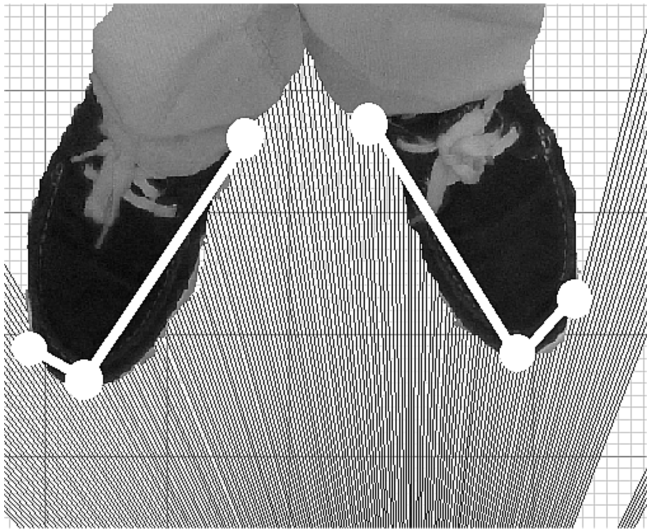

Figure 4. Detected data and extracted feature points.

Data

We collected data over the course of one calendar week, from Monday to Friday, 9 am to 7 pm. We recorded video, touch-input, and foot pattern data and collected total foot patterns of 115 from 84 users. Data from a group of users was sorted out. A foot pattern was logged automatically, when being static and at a maximal distance of 1.5 m from the screen. The average duration of one recorded foot pattern was 9 s and the average duration of one interaction 11 s. If the user changed her posture during interaction, the new pattern was logged as new data in the same intention category. We distinguished between users who need information and users who don't need information using Norman's execution-evaluation cycle [14]. To simplify the model, the user's understanding of the interface's content and functionality is left out of this study. The execution state while touching the display was followed by an evaluation of the system state with respect to the user's intention.

Intention Categories

Using the video recordings and touch-log, we manually classified user behavior after the touching of the display. We assume that the activity of the user is directly related to her intention. If the evaluation phase was followed by a click of one of the other menu buttons, the user's intention differed from the current system state. We then concluded that the system had been used out of curiosity. We refer to this intention as Intention B. If the user had the goal of accessing information, he/she remained in the interpretation state. We referred to this state as Intention A, or interaction to access specific information. If the sample seemed ambiguous, we referred to the displayed content: If the accessed content was the map, we categorized intention A, otherwise we categorized B. Another category was, if the user also asked staff for directions before or after interaction.

Based on these criteria, we placed each user in either category A or B. If the user changed her posture, we added the new foot pattern to the dataset accordingly. 48% of the 115 samples were indicated as intention A, which means half of the users were using the system to access specific information. This gives us a dataset of two separated intentions and captured foot patterns.

Discussion

Based on the literature regarding *foot language*, the direction of the feet indicates the spatial interest of a person [2, 3]. To prove whether this is applicable at public displays, we analyzed the foot pattern angles and positions. Figure 5 shows sample foot patterns.

Foot patterns similar to those samples were captured frequently. Foot patterns occur as described in [2, 3]: the foot, as an indicator for walking direction and stable platforms, expresses intentions related to agents we interact with. If the user requests specific information, she will posture herself stable and orientated to the screen. If she does not need specific information and uses the system out of curiosity, the feet are pointing away with intention to leave.

Figure 5. Sample foot patterns that frequently occurred.

	Predict A	Predict B
Actually A	92.0% (23)	8.0% (2)
Actually B	23.3% (7)	76.7% (23)
	Accuracy: 84.4%	

Figure 6. Results from SVM algorithm.

As a final step we classified the foot data using the SVM algorithm in [15]. We used the feature data (two foot angles and the position) of 30 recorded samples from each intention as training data and tested the algorithm using the 25 residual samples from Intention A and 30 samples from Intention B. Figure 6 shows the resulting confusion matrix with an accuracy of 84.4%, which is significant and shows that users position their feet according to spatial interest. The proposed framework might be further developed through integration of additional modalities.

Conclusion

In this note, we demonstrated that foot patterns make a significant contribution in the measurement of user intention in public display environments. Using a laser rangefinder, a touch-screen log, and video data, we studied users at a public multi-touch display. We separated user intention into users who need information and users who don't need information. We show that each of these user groups had different, characteristic foot patterns.

Integrating foot measurement systems with public displays is an approach to enhancing human-computer dialogue in such displays. Information regarding user intention is valuable for studies seeking public displays with increased interaction, e.g. adaptive display content without asking the user explicit questions. Approaches applying interaction proxemics can be expanded using foot pattern expressions. Although it seems clear that foot patterns vary from user to user, we expect that it can improve the awareness of the system, or at least support complex sensing systems. As a next step, the intention classifications could be improved by including interviews which could also improve accuracy results.

In the end, foot pattern recognition could enhance a variety of HCI applications in public display research.

The reliability of an intention detection system could be increased with the combination of foot tracking and face tracking. Foot tracking will be further improved by experiments in a variety of locations. Also it would be interesting to explore whether the proposed technique applies in other contexts, e.g. when sitting or during interaction with a robot. Furthermore, non-average foot positions, e.g. from disabled people, require more in-depth study. All these applications will be important steps toward the ultimate goal of a system that understands nonverbal expressions just as humans do.

Acknowledgements

This work was made possible by Dr. Ji-Hyung Park's mentorship and supported by a grant from the KIST Institutional Program.

References

[1] Hall, E.T., The hidden dimension. Vol. 6., 1966.

[2] Morris, D., Bodywatching: A field guide to the human species. 1985: Cape.

[3] Morris, D., Peoplewatching: The Desmond Morris Guide to Body Language. 2002: Vintage Books.

[4] Michelis, D. and J. Müller, The audience funnel: Observations of gesture based interaction with multiple large displays in a city center. Intl. Journal of Human–Computer Interaction, 2011. 27(6): p. 562-579.

[5] Vogel D., Balakrishnan R. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. UIST 2004, ACM.

[6] Peltonen, P., E. Kurvinen, A. Salovaara, G. Jaccucci, T. Ilmonen, J. Evans, A. Oulasvirta, P. Saarikko, It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. 2008. ACM CHI.

[7] Brignull, H. and Y. Rogers, Enticing people to interact with large public displays in public spaces. Proceedings of INTERACT, 2003.

[8] Nakanishi, Y., T. Fujii, K. Kiatjima, Y. Sato, H. Koike, Vision-based face tracking system for large displays. UbiComp 2002: Ubiquitous Computing, 2002: p. 123-160.

[9] Vending Machine Uses Facial Recognition To Make Recommendations Based On Stereotypes http://www.geekologie.com/2011/12/vending-machine-uses-facial-recognition.php

[10] Annett, M., T. Grossman, D. Wigdor, G. Fitzmaurice, Medusa: a proximity-aware multi-touch tabletop. 2011. UIST 2011, ACM.

[11] Augsten, T., K. Kaefer, R. Meusel, C. Fetzer, D. Kanitz, T. Stoff, T. Becker, C. Holz, P. Baudisch, Multitoe: high-precision interaction with back-projected floors based on high-resolution multi-touch input. UIST 2010, ACM.

[12] Huan, E., A. Koster, and J. Borchers, Overcoming assumptions and uncovering practices: When does the public really look at public displays? Pervasive Computing, 2008: p. 228-243.

[13] Arras, K.O., S. Grzonska, M. Luber, W. Burgard, Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities. ICRA 2008, IEEE.

[14] Dix, A., Human-computer interaction. 2004: Prentice hall.

[15] Chang C., Lin C., LIBSVM -- A Library for Support Vector Machines. http://www.csie.ntu.edu.tw/~cjlin/libs