Cross-Surface

Challenges and Opportunities of Spatial and Proxemic Interaction

2016-2

Proceedings of the third Cross-Surface workshop

Editors: Steven Houben, Jo Vermeulen, Nicolai Marquardt, Johannes Schöning, Clemens Klokmose and Harald Reiterer.

Publication: Published online in 2016.

Contact: **Steven Houben (s.houben@lancaster.ac.uk)**

About

In this workshop, we reviewed and discussed open issues, technical challenges and conceptual models for multi-device spatial or proxemic interaction. We brought together researchers, students and practitioners working on technical infrastructures, studies and designs of spatial interfaces, or domain specific multi-device applications that use space as a unit of analysis. We focused specifically on analyzing how such interfaces, tools and tracking technology can be deployed "in the wild". Please find more details about the workshop, in the submitted proposal [1]. The workshop was held in conjunction with the 2016 ACM International Conference on Interactive Surfaces and Spaces (ISS), that took place from November 6 to 9 in Niagara Falls, Canada.

[1] Houben, Steven, Jo Vermeulen, Clemens Klokmose, Johannes Schöning, Nicolai Marquardt, and Harald Reiterer. "Cross-surface: Challenges and opportunities of spatial and proxemic interaction." In Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces, pp. 509-512. ACM, 2016.
DOI: https://doi.org/10.1145/2992154.2996360

Editors

Dr. Steven Houben – *Lancaster University*
Dr. Nicolai Marquardt – *University College London*
Dr. Jo Vermeulen – *University of Calgary*
Prof. Johannes Schöning – *Hasselt University – tUL –iMinds*
Prof. Clemens Klokmose – *Aarhus University*
Prof. Harald Reiterer – *University of Konstanz*

Program

08:00 Registration
09:00 Tutorial: "Cross-Device Interfaces: Existing Research, Current Tools, Outlook"
10:30 Coffee Break
11.00 *Topic 1*: Combining Devices
11:45 *Topic 2*: Devices and Collaboration
12:30 Lunch
13:30 Interactive sessions
15:00 Coffee break
15:30 *Topic 3*: Spatial Interactions
16:30 Closing
18:00 ISS Welcome Reception

Keynote – Tutorial

Title: “Cross-Device Interfaces: Existing Research, Current Tools, Outlook”

In this tutorial, I will describe key technical and design challenges of developing cross-device interfaces, which are not only able to adapt to a wide variety of device characteristics and user preferences, but can actually migrate a user’s tasks from one device to another and even distribute the interface between multiple devices jointly used by one or more users. Starting with a review of the cross-device design concepts, authoring tools and interaction techniques presented in the ISS and larger HCI literature, I will give a comprehensive overview of the state of the art. I will then analyze the current tools for design and development of cross-device interfaces, and present one of these tools, XDBrowser, in detail. Participants will get hands-on experience with designing cross-device versions of existing interfaces using such authoring tools created for non-technical users. Finally, I will discuss open issues and provide an outlook based on the results from recent cross-device workshops I co-organized or participated in at CHI, EICS, and DIS.

Biography: *dr. Michael Nebeling is an Assistant Professor at the University of Michigan School of Information, where he leads the Information Interaction Lab. In this lab, they study the next generation of user interfaces, as well as the methods and tools to create them. With the goal of making interfaces easier to use and to create, their research benefits both users and designers, and allows more users to be designers.*

List of accepted papers

1. **Combining Physical and Social Proximity for Device Pairing**
   Maria Husmann, Sivaranjini Chithambaram and Moira Norrie
   ETH Zurich
2. **Discussing the State of the Art for "in the wild" Mobile Device Localization**
   Tom Horak, Ulrich von Zadow, Matthias Kalms and Raimund Dachselt
   Technische Universität Dresden
3. **A Proposal for a Taxonomy to Survey the State-of-the-Art of Multi-Device Technologies**
   Roman Rädle [a], Clemens N. Klokmose [a] and Hans-Christian Jetter [b]
   [a] Aarhus University
   [b] University of Applied Sciences Upper Austria
4. **Interactions in a Human-Scale Immersive Environment: the CRAIVE-Lab**
   Gyanendra Sharma, Jonas Braasch and Richard Radke
   Rensselaer Polytechnic Institute
5. **A Tool for Multi-Surface Collaborative Sketching**
   Jorge Luis Pérez Medina and Jean Vanderdonckt
   Université catholique de Louvain, Louvain Interaction Lab
6. **Device Orientation Sensors for Pairing and Spatial Awareness in the Wild**
   Jens Emil Grønbæk [a] and Kenton O'Hara [b]
   [a] Aarhus University
   [b] Microsoft Research
7. **Using Low Cost Multi-Sensory Output Cues to Support Proxemics and Kinesics Across Heterogeneous Systems**
   Rajiv Khadka [ab], James Money [b] and Amy Banic [ab]
   [a] University of Wyoming
   [b] Idaho National Laboratory
8. **Designing Multi-Surface Environments to Support Collaborative Sensemaking**
   Leila Homaeian, James Wallace and Stacey Scott
   University of Waterloo
9. **Considering Collaboration in ʔeləw̓k̓ʷ — Belonging**
   Reese Muntean
   Simon Fraser University
10. **Towards Interaction Around Unmodified Camera-equipped Mobile Devices**
    Jens Grubert [a], Eyal Ofek [b], Michel Pahud [b], Matthias Kranz [c], and Dieter Schmalstieg [d]
    [a] Coburg University
    [b] Microsoft Research
    [c] University of Passau
    [d] Graz University of Technology
11. **Just Scratching the Surface, the Long Road to Effective Cross-Display Interaction**
    Narges Mahyar [a], Kellogg Booth [b], Cynthia Girling [b] and Ronald Kellett [b]
    [a] University of California
    [b] University of British Columbia

Combining Physical and Social Proximity for Device Pairing

Maria Husmann
Department of Computer Science
ETH Zurich
husmann@inf.ethz.ch

Sivaranjini Chithambaram
Department of Computer Science
ETH Zurich
chithams@student.ethz.ch

Moira C. Norrie
Department of Computer Science
ETH Zurich
norrie@inf.ethz.ch

Presented at the Cross-Surface '16 workshop, in conjunction with ACM ISS'16.
November 6, Niagara Falls, Canada.

Abstract
Cross-device interaction often starts with a pairing step where devices are associated. Pairing that is based purely on physical proximity risks accidental or even hostile pairing when users are physically but not socially close. In this paper, we explore how data on a user's social circle can inform the pairing process. We propose three pairing modes that balance user control with ease of use and trigger them based on social and physical proximity of the devices and users involved.

Author Keywords
cross-device; social proximity; physical proximity; device pairing

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction
As the number of devices in our lives has increased, researchers have been investigating how these devices can be used in combination [5, 10, 8]. Both single and multi-user scenarios have been explored [4]. To enable the usage of multiple devices in a coordinated manner, there needs to be an association step connecting the devices. A number of different approaches and interaction techniques have

been explored (see [3] for a summary). One option is to rely purely on proximity and allow interaction among all devices in a certain range, for example the same Wifi network or in range of a Bluetooth connection. However, these approaches come with an inherent security risk as proximity does not necessarily imply an interest in interacting, as illustrated by the case of a woman who received indecent images from a nearby passenger on a train [2]. The woman had enabled Airdrop to share a photo with someone else which allowed the perpetrator to also send pictures to her. There are security settings that allow the user to restrict access to people in the contact list, however, the user has to remember to switch to the right settings depending on who they want to share content with.

Another option is to require some manual coordination, for example a synchronous gesture with physical contact such as bumping two devices together [6], or continuous gestures that start on one device and continue on the next, for example stitching [7].

We propose another approach for pairing devices that adds *social proximity* as a dimension. Our main idea is to make device pairing very easy for devices with the same or socially close owners, and gradually add more steps as social distance increases.

Scenarios

Physical proximity does not necessarily imply social proximity, but rather is an orthogonal dimension. Figure 1 illustrates these two dimensions with sample scenarios.

1. In the top left corner, we have people who are physically but not socially close to each other, such as strangers on a train.
2. When we move to the top right, we decrease social

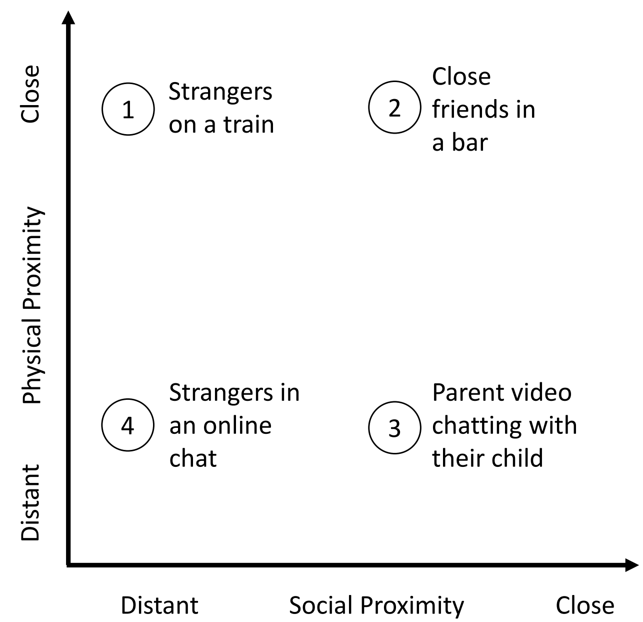

Figure 1: Scenarios of social and physical proximity.

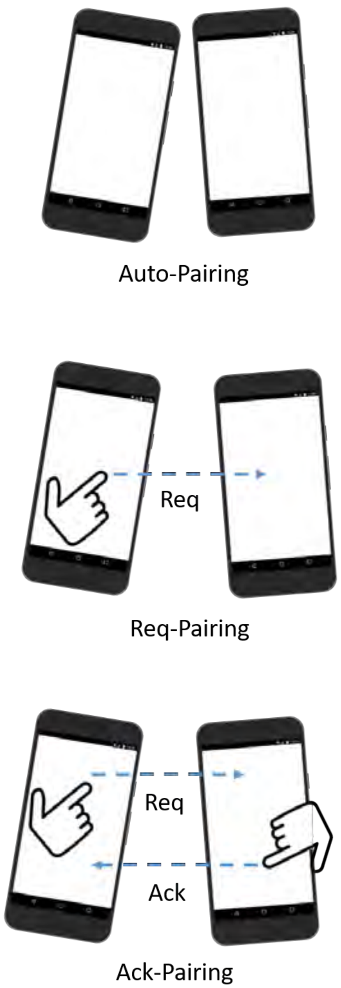

Figure 2: The three pairing modes with varying user control.

distance. An example here would be close friends spending time together in a bar.

3. In the bottom right, we have people who are socially close but not physically, such as a parent video chatting with their child.

4. Situated in the bottom left corner are people who are neither physically nor socially close, for example strangers chatting on the internet.

Our main interest is to provide a pairing mechanism for scenarios 2 and 3, while avoiding accidental pairing in scenario 1. Scenario 4 is not considered here since existing methods such as sharing a pairing code over an alternative channel such as a chat are sufficient.

Social Circles

Our social contacts can be grouped into circles [9], for example family, close friends, bike buddies or colleagues. Social networks such as Facebook or Google+ reflect this structure in their user interface and provide mechanisms such as friend lists or circles to organise one's contacts. Facebook even tries to automate this process to some extent with their Smart Lists[1] that group friends based on profile information, such as working for the same company. Furthermore, there are two standard lists, *close friends* and *acquaintances*. On the one hand, these lists can be used to tailor the information in a user's news feed to show more updates from close friends and fewer from more distant acquaintances. On the other hand, this information allows a user to share content with only a specific subset of contacts, for example all friends except acquaintances or only family.

[1] https://www.facebook.com/notes/facebook/improved-friend-lists/10150278932602131/

We propose that this information could be used to determine social distance, which, in turn, could help select an optimal pairing mode. The closer the owner of a given device is to us socially, the easier it should be to pair our devices. For more socially distant people, such as acquaintances, more control should be given to each of the users involved in the pairing.

Pairing Modes

We propose three different pairing modes (Fig. 2) for people who are connected via a social network.

- *Auto-Pairing* automatically pairs two devices and requires no user interaction.
- *Req-Pairing* requires a pairing request from the initiating device. No interaction is required from the receiving device.
- *Ack-Pairing* requires a pairing request from the initiating device and the receiver has to accept the request in order to establish the connection.

The modes illustrate the trade-off between ease of use and control. While the first mode (auto) requires no user interaction, it also offers no direct control over the pairing. In contrast, the last mode (ack) gives control to both involved users, however, it also requires both to explicitly interact with device. We propose that the appropriate mode should be chosen based on both the social and physical proximity of the involved users. Devices belonging to the same user are considered to have high social proximity and should be paired automatically if they are sufficiently close physically. As physical distance increases, Req-Pairing should be used. The assumption here is that a user would frequently pair devices that are physically close, for example in the

Figure 3: Contacts and their devices (represented by an icon) that are available for pairing, showing closest devices first.

same room, and this should be made as easy as possible. At the same time, this approach avoids that devices that are distant are paired accidentally, such as a user's phone in the office being paired with their computer at home. There could be occasions where this is desired, but we assume that these are rare, and would be supported with an explicit pairing request from the user. For devices that belong to socially close people, such as close friends or family, we propose Req-Pairing and, for socially more distant people, such as colleagues or acquaintances, Ack-Pairing could be used. Physical distance could be used to only display contacts within a specific distance or to sort contacts by distance, showing closer ones first.

For people who are physically close, but who do not (yet) have a social connection, we propose that existing proximity-based approaches such as scanning a QR code could be used for pairing.

Prototype Implementation
We have implemented the pairing modes outlined above in a prototype as an extension to our web-based cross-device framework XD-MVC[2] and used them in two sample applications. One challenge in the implementation was accessing a user's social circles and we have not yet found a satisfactory solution. While this information is available in Facebook and Google+, their APIs do not allow third-party applications to access it for privacy reasons. Google+ offers access to a user's circle via the Domains API[3], however, this API is only available to paid, corporate Google Apps accounts. Our workaround is based on Google's People API[4] and relationships are determined with the *relationship* attribute in a user's contacts. However, we assume that this attribute is rarely used and access to Facebook's friend lists would yield better results. Physical distance was determined based on the Geolocation API[5] that is provided by modern browsers.

Let us illustrate how the pairing modes can be used from a user perspective based on our sample application *Shared-Booking* which is a hotel booking application that distributes its user interface among paired devices (Fig. 4). As a first step, the application needs to be opened on all devices that are to be paired and the user needs to authenticate. The authentication can be stored on the device, so this only needs to be done the first time the application is used. If the same user is authenticated on multiple devices, these pair automatically and trigger redistribution of the user interface. In addition, there is a component in the UI that lists a user's contacts who also have the application open at the same time (Fig. 3). By default, the list is sorted to show physically close devices first. When the user selects a device, either *Req-* (for close friends and family) or *Ack-Pairing* is used. In the latter case, a popup is shown on the selected device asking its owner to confirm the pairing.

Since social relationships are not necessarily reciprocal [1], we choose the pairing mode based on the more distant relationship.

Discussion
Our prototype implementation has demonstrated that there are still a lot of open issues and questions. First, access to social circles is not easy to obtain with current API, despite the existence of the information in current social networks.

The fact that the user needs to open their application and

[2] https://github.com/mhusm/XD-MVC
[3] https://developers.google.com/+/domains/getting-started
[4] https://developers.google.com/people/
[5] https://developer.mozilla.org/en-US/docs/Web/API/Geolocation/Using_geolocation

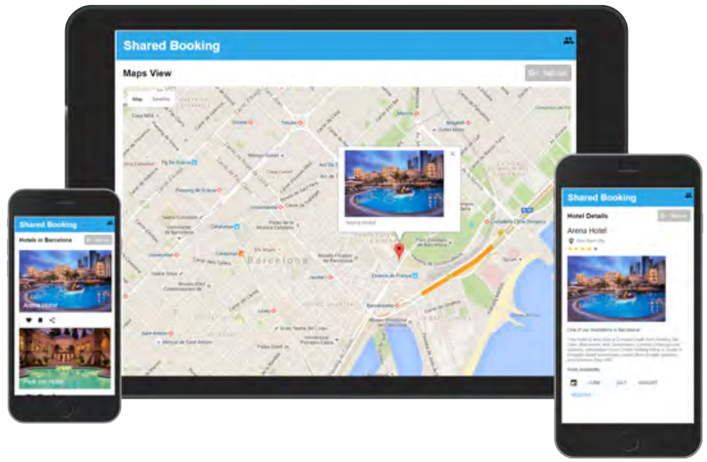

Figure 4: The SharedBooking sample application when used on three paired devices. The left device shows a list of hotels, the tablet in the middle a map marking the hotel chosen from the list, and the right phone shows details of the chosen hotel.

Figure 5: In *Ack-Pairing* mode a popup asks the user to confirm the pairing request.

be authenticated before the actual pairing implies that this approach is best suited for applications that are used frequently and accessible easily from a devic, for example installed on the home screen. Other approaches, such as pairing via QR codes or NFC allow the application URL to be transmitted along with a pairing token and may be better suited for less frequently used applications.

Our approach is based on the notion that people maintain lists of friends that reflect social distance. We have not confirmed to what extent this is actually the case. Alternatively, a social network provider could attempt to estimate the distance based on communication patterns. Furthermore, we have not yet validated our pairing modes with users. A user study could be conducted to investigate, if they meet actual user needs and to fine tune the social and physical boundaries that trigger each mode.

Finally, information on social circles could be used in cross-device applications for purposes other than pairing. Assume two users pair their mobile devices for photo sharing. After the pairing has been established, information on social proximity could be used to determine what pictures should be accessible on each device.

Conclusion

We have investigated how physical and social proximity can be combined to facilitate the pairing of devices and avoid accidental pairing that can occur when pairing is done solely based physical proximity. We have proposed three pairing modes and implemented them in a prototype. Our preliminary implementation showed that current social network APIs are lacking and that further research into using social network data in cross-device applications is needed.

Acknowledgements

This project was supported by grant No. 150189 of the Swiss National Science Foundation (SNF). Discussion during the first Cross-Surface workshop in 2015 stimulated this work.

REFERENCES

1. Abdullah Almaatouq, Laura Radaelli, Alex Pentland, and Erez Shmueli. 2016. Are You Your Friends' Friend? Poor Perception of Friendship Ties Limits the Ability to Promote Behavioral Change. *PloS one* 11, 3 (2016).

2. Sarah Bell. 2015. Police investigate 'first cyber-flashing' case. BBC News. (13 August 2015). Retrieved September 14, 2016 from http://www.bbc.com/news/technology-33889225.

3. Ming Ki Chong, Rene Mayrhofer, and Hans Gellersen. 2014. A Survey of User Interaction for Spontaneous Device Association. *ACM Comput. Surv.* 47, 1 (May 2014). DOI:http://dx.doi.org/10.1145/2597768

4. Linda Di Geronimo, Maria Husmann, and Moira C. Norrie. 2016. Surveying Personal Device Ecosystems with Cross-device Applications in Mind. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays (PerDis '16)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/2914920.2915028

5. Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: Enabling and Understanding Cross-Device Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2773–2782. http://doi.acm.org/10.1145/2556288.2557170

6. Ken Hinckley. 2003. Synchronous Gestures for Multiple Persons and Computers. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology (UIST '03)*. ACM, New York, NY, USA. DOI:http://dx.doi.org/10.1145/964696.964713

7. Ken Hinckley, Gonzalo Ramos, Francois Guimbretiere, Patrick Baudisch, and Marc Smith. 2004. Stitching: Pen Gestures That Span Multiple Displays. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '04)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/989863.989866

8. Steven Houben, Paolo Tell, and Jakob E. Bardram. 2014. ActivitySpace: Managing Device Ecologies in an Activity-Centric Configuration Space. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/2669485.2669493

9. Charles Kadushin. 1966. The Friends and Supporters of Psychotherapy: on Social Circles in Urban Life. *American Sociological Review* (1966), 786–802.

10. Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. ACM, New York, NY, USA, 45–54. DOI: http://dx.doi.org/10.1145/2669485.2669500

Discussing the State of the Art for “in the wild” Mobile Device Localization

Tom Horak
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
horakt@acm.org

Raimund Dachselt
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
dachselt@acm.org

Ulrich von Zadow
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
uzadow@acm.org

Matthias Kalms
Interactive Media Lab
Technische Universität Dresden
Dresden, Germany
matthias.kalms@mailbox.tu-dresden.de

Presented at the *Cross-Surface ’16 workshop*, in conjunction with ACM ISS’16.
November 6, Niagara Falls, Canada.

Abstract
Technologies for spatial and proxemic interaction with mobile devices depend inherently on the ability to obtain information on the device’s position (i.e., to *localize* the device). Numerous technologies have been proposed for this, each with their own strengths and weaknesses, but deciding which one to use in a particular context is challenging. In this paper, we examine current technologies for the localization of mobile devices and categorize them into a taxonomy based on their technological similarity. By considering numerous properties (e.g., precision, battery usage, scalability, required infrastructure, deployment) and discussing how these impact usability in different scenarios, we aim to allow other researchers informed decisions on the localization techniques to use for a particular application case.

Author Keywords
Localization Techniques; in the wild; Mobile Devices

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

Introduction
Research on spatial interaction with mobile devices has mostly used instrumented devices and rooms such as marker-based tracking [2, 11, 32] (or, in the early days, teth-

Signal Spectrum	radio-, light-, or sound-based
Technical Base	hardware or software base used
Computation Method	position determination method
Localization Approach	positioning or tracking
Precision	metric precision based on literature, ranging from mm to >m
Range	estimated metric range, ranging from km to m
Battery Usage	estimated usage, ranging from none to high
Scalability	scalibility of number of devices, limited or ∞ (unlimited)
Required Infrastructure	infrastructure used, rated by installation effort
Device Instrumentation	required device instrumentation
Deployment Effort	estimated effort of deployment process; none to high
Deployment Costs	estimated costs of deployment; none to high
Examples	related literature
Challenges	existing/important challenges

Table 1: Properties and their possible values for Table 2

ered devices [30]) to determine devices' positions. While this approach enables research on interaction, deploying the developed techniques in the wild has additional requirements: Among others, instrumentation of the mobile device will not be possible in most cases, mobile power consumption and scalability become an issue, and deployment efforts as well as costs need to be considered. Further, requirements often differ depending on the application case. As an example, while many single-device outdoor applications (e.g., wayfinding) work well with coarse positioning, indoor applications (e.g., pointing for data transfer methods [11, 32]) rely on much more precise positioning, while the range of the technique can be as low as a few meters.

In this position paper, we first examine important properties for "in the wild" localization techniques (see Table 1). Based on these properties, we then categorize existing localization techniques into a taxonomy and present these in a tabular form (Table 2). We color-coded the properties in the table to allow to quickly recognize advantages and challenges of the different techniques. Finally, we discuss the requirements of different application cases, with the goal to enable informed decisions on the techniques to use.

Properties of Localization Techniques

First of all, most localization techniques are based on the usage of signal waves and differ in the **signal spectrum** used (e.g., radio, light, sound). This spectrum heavily influences other properties and defines also the main source of disturbance—other waves in the same spectrum (e.g., sun light). All techniques build upon a certain **technical base** and use a specific **computation method**. Depending on the latter one, the **localization approach** can be either a positioning or a tracking approach, i.e., the position can be calculated by the device itself or by the infrastructure. We consider three main performance properties resulting from the signal, technical base and computation method used: The **precision** ranging from millimeters to multiple meters, the signal **range** from a few meters to many kilometers, and the extra device **battery usage** from none to high.

Especially important for "in the wild" application scenarios are properties influencing the deployment of spatial tracking. Regarding the hardware, this involves the **scalability** (i.e., if more devices can be easily incorporated), the **required infrastructure** (i.e., none, existing, or additional), as well as the the need for **additional instrumentation** for consumer devices used. Regarding the scalability, we only differentiate between a (practical) unlimited and a limited scalability (e.g., limited by increasing synchronization problems). Based on the required deployment steps and hardware, we roughly rated all techniques regarding their relative **deployment effort** and **deployment costs** compared to other techniques from none to high. These deployment properties can only serve as rough indicator. Finally, we also list **examples** from related work and name existing main **challenges** of the localization techniques.

Localization Techniques

For our taxonomy, we use the signal spectrum as main category and distinguish between three groups: radio, light, and (ultra-)sound (top level of taxonomy). In each group, we further differentiate between which technical base (2nd level) and computation method (3rd level) is used. We characterize the resulting localization techniques based on the properties described before (see also property overview in Table 1). Furthermore, we color-coded—and thus rated—all property values in three steps (green, yellow, red) roughly indicating their usability. We do not consider techniques based on inertial sensors only (e.g., gyroscope), as they lack the ability of localization relative to other devices. The complete taxonomy is shown in Table 2 at page 3.

Signal Spectrum	Technical Base	Computation Method	Localization Approach	Precision	Range	Battery Usage	Scalability	Required Infrastructure	Device Instrumentation	Deployment Effort	Deployment Costs	Examples	Challenges
radio	GPS	TDOA	positioning	> m	km	low	∞	existing	none	none	none	car navigation	outdoor only
	GSM	RSSI	positioning	> m	km	low	∞	existing	none	none	none	-	mainly in urban regions
		fingerprinting	positioning	m	km	low	∞	existing	none	high	none	[3, 20]	mainly in urban regions, reference model
	WiFi	RSSI	positioning	m	> m	low	∞	existing	none	none	low	[1, 6]	-
		fingerprinting	positioning	< m	> m	low	∞	existing	none	high	low	[1, 3, 16, 35]	reference model
		TOA/TDOA	positioning	m	> m	mid	∞	existing	none	none	low	[34]	multipath effects
		AOA	positioning	< m	m	mid	∞	existing	none	none	low	[12, 33]	multipath effects
	Bluetooth	RSSI	positioning	cm	m	low	limited	none	none	mid	none	[9]	adaptive signale strength
		fingerprinting	positioning	cm	m	low	limited	none	none	high	low	[3, 10]	adpative signale strength, reference model
light	Device RGB Cam	triangulation	positioning	cm	m	high	∞	markers	none	low	low	[15]	limited field of view, marker placement
		RSSI	positioning	m	> m	mid	∞	landmarks	none	high	low	[24]	number and type of required landmarks
		feat. detection	positioning	cm	> m	high	∞	none	none	none	none	[4], Proj. Tango	calibration / training
	External RGB Cam	triangulation	tracking	mm	> m	none	limited	camera	markers	mid	mid	-	limited field of view
		rect. detection	tracking	cm	> m	none	limited	camera	none	low	low	[26]	disturbance from overhead lights
	External IR Cam	triangulation	tracking	mm	> m	none	limited	cameras	markers	mid	high	OptiTrack, Vicon	calibration
		depth (TOF)	tracking	cm	> m	none	limited	camera	none	low	mid	[26, 27]	limited field of view
(ultra-)sound	Audio Devices	fingerprinting	positioning	m	> m	mid	∞	none	none	high	none	[25, 28, 29]	reference model
		TOA	tracking	cm	> m	low	limited	micros	none	mid	mid	[18, 22]	synchronization, multipath effects
		TOA	positioning	cm	> m	mid	∞	speaker	none	mid	low	[13, 19]	synchronization, multipath effects
		RT-TOA	positioning	cm	> m	mid	limited	none	none	none	none	[9, 16, 21, 23]	range internal hardware, multipath effects
		TDOA	tracking	cm	> m	low	limited	micros	none	mid	mid	[7]	multipath effects
		TDOA	positioning	cm	> m	mid	∞	speaker	none	mid	low	[13, 14, 17]	multipath effects

Table 2: Localization techniques and their color-coded properties grouped by signal spectrum and their technical base.

Radio-based
Many existing systems use radio frequency signals for localization, as current commodity hardware support them by default. Most prominently, GPS provides reliable localization at a precision of multiple meters in outdoor scenarios. For indoor scenarios, GSM base stations, WiFi access points and Bluetooth devices can serve as a technical base for localization providing up to meter, sub-meter or centimeter precision respectively. Received signal strength indicators (RSSI) can both be used to directly infer distance [1, 6, 9] or be compared as fingerprint to a prerecorded model of the signal in an area [1, 3, 10, 16, 20, 35]. Techniques leveraging the angle of arrival (AOA) [12, 33], time of arrival (TOA) or time difference of arrival (TDOA) [34] for multiple sources have also been explored.

Light-based
Optical techniques for localization typically incorporate either standard video cameras or infrared (IR) cameras and can be marker-based or not. Further, they can be distinguished into techniques either using an internal camera (i.e., device camera) or external cameras, thus between positioning and tracking systems respectively. For the latter one, professional setups like OptiTrack[1] or Vicon[2] usually track IR-reflecting markers, thus requiring instrumentation of the devices as well as the environment (i.e., placing cameras). Recently, low-cost—but less precise—depth-camera-based systems were presented [26, 27]. These are markerless and thus can be more easily deployed.

Positioning systems using an internal device camera come with the cost of higher battery drain, but often also incorporate a reduced deployment effort. The device camera detects prior installed reference points, e.g., light landmarks [24], markers [15], or image patterns [4], and the device can calculate its position based on these points. With Google's Project Tango[3] there also exists a solution that does not require any installed reference points but detect features of the environment by its own. However, to allow an absolute localization, an initial calibration step is required.

(Ultra-)sound-based
Sound-based techniques are an interesting alternative to the well-established radio- and light-based localization providing centimeter precision at low deployment efforts. Localization can be calculated by determining the distance to reference points from TOA [13, 18, 19, 22] or TDOA [7, 13, 14, 16, 17]. Due to the relatively slow propagation of sound, even standard speakers and microphones achieve sufficient measurement accuracy. Therefore, based on the same techniques both positioning [13, 14] and tracking [7, 18, 22] systems are possible. Using round-trip time of arrival (RT-TOA) [9, 21, 23] localization can also be performed among devices without any external infrastructure, however, providing only relative positioning in this case (RT-TOA can also be used with fixed infrastructure). Similar to RSSI-based fingerprinting, sound profiles of ambient noise [25] and purposely installed sound sources [28] can be utilized for positioning at room level. By recording the reflections of sound signals emitted by the device itself, even centimeter precision is possible [29]. However, changes in temperature and humidity affect sound propagation requiring to take environmental changes into account.

Improving Localization with Sensor Fusion
Combining different localization techniques with complementary attributes can help to improve overall performance. A common scenario is to pair a reliable but less accurate

[1] https://www.optitrack.com/
[2] https://www.vicon.com
[3] https://www.google.com/atap/project-tango/

technique with a less reliable, accurate one (e.g., [3, 9, 13, 16, 26]). Choosing techniques from different signal categories not only evades mutual interference, but also allows to compensate interference in one of them. Further, also internal device sensors (e.g., gyroscope) can be used to enhance localization with fine-grained movement and rotation detection (e.g., Project Tango or [9]).

Discussion & Conclusion

The required properties of localization techniques depend heavily on the application case (e.g., home, office, indoor public spaces, outdoor). For instance, in the context of private or home applications (e.g., [2]) the localization must come at low costs, but can involve instrumentation of rooms or devices and must not support a large number of devices. In offices, the number of devices is still manageable, but device instrumentation is probably not appropriate. One specific application scenario are smart meeting rooms, which might support pointing interactions [4, 5, 8], thus resulting in the highest demand on precision. Both personal and business application cases deploy localization techniques in controlled environments with minor sources of disturbance and small number of devices.

In contrast, indoor public spaces (e.g., shopping mall, museum) are crowded places in which localization techniques must be able to handle many devices at the same time in a larger area. These places often incorporate many barriers (e.g., walls, people) and manifold sources of disturbance, especially for light-based localization techniques (e.g., light installations, reflecting surfaces). A common application case is enabling cross-device interactions between a larger display and personal devices, e.g., to view, explore, or transfer information [4, 31]. In shopping malls, the instrumentation of devices is not feasible and additional infrastructure may have to fulfill some aesthetic requirements, whereas in museums specialized devices can be handed out and visible cameras are more accepted. As example environments for "in the wild" localization, both indoor and outdoor public spaces usually also have to support a wider variety on devices, thus rely more on commodity solutions than personal or business scenarios.

Taking this further, an ideal localization technique would be characterized by requiring only existing consumer hardware, a minimal deployment effort, an unlimited scalability, few sources of disturbance, a low battery usage, and, of course, a high precision. This mainly rules out light-based techniques, as they often require a comprehensive instrumentation or drain the battery. In contrast, radio-based techniques can often be easily deployed as they rely on existing hardware. However, the precision is poorer as for light-based techniques. Sound-based techniques could evolve as an alternative between these two established ones as they offer up to centimeter precision with reasonable deployment effort. Deploying "in the wild" localization systems in a larger scale also raises questions if and how this could affect our society, e.g., regarding privacy concerns (tracking vs. positioning; who has access to tracking data) or environmental aspects (such as noise pollution).

Still, choosing the *right* localization technique depends on the specific application and the resulting interaction style. We are confident that our discussion and presented taxonomy in Table 2 can help to identify the most suitable candidates and thus supports the process of deciding which localization technique to use for a specific application case.

REFERENCES

1. Bahl, P., and Padmanabhan, V. N. Radar: An in-building rf-based user location and tracking system. In *Proc. INFOCOM '00*, vol. 2, IEEE (2000), 775–784.

2. Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic interaction: Designing for a proximity and orientation-aware environment. In *Proc. ITS '10*, ACM (2010), 121–130.

3. Bolliger, P. Redpin - adaptive, zero-configuration indoor localization through user collaboration. In *Proc. MELT '08*, ACM (2008), 55–60.

4. Boring, S., Baur, D., Butz, A., Gustafson, S., and Baudisch, P. Touch projector: Mobile interaction through video. In *Proc. CHI '10*, ACM (2010), 2287–2296.

5. Bragdon, A., DeLine, R., Hinckley, K., and Morris, M. R. Code space: Touch + air gesture hybrid interactions for supporting developer meetings. In *Proc. ITS '11*, ACM (2011), 212–221.

6. Chintalapudi, K., Padmanabha Iyer, A., and Padmanabhan, V. N. Indoor localization without the pain. In *Proc. MobiCom '10*, ACM (2010), 173–184.

7. Höflinger, F., Zhang, R., Hoppe, J., Bannoura, A., Reindl, L. M., Wendeberg, J., Bührer, M., and Schindelhauer, C. Acoustic self-calibrating system for indoor smartphone tracking (assist). In *Proc. IPIN '12* (2012), 1–9.

8. Horak, T., Kister, U., and Dachselt, R. Presenting business data: Challenges during board meetings in multi-display environments. In *Proc. ISS '16*, ACM (2016).

9. Jin, H., Holz, C., and Hornbæk, K. Tracko: Ad-hoc mobile 3d tracking using bluetooth low energy and inaudible signals for cross-device interaction. In *Proc. UIST '15*, ACM (2015), 147–156.

10. Jin, H., Xu, C., and Lyons, K. Corona: Positioning adjacent device with asymmetric bluetooth low energy rssi distributions. In *Proc. UIST '15*, ACM (2015), 175–179.

11. Kister, U., Klamka, K., and Dachselt, R. Supporting graph exploration tasks on display walls using spatially-aware mobile devices. In *Proc. InfoVis '16*, IEEE (2016).

12. Kotaru, M., Joshi, K., Bharadia, D., and Katti, S. Spotfi: Decimeter level localization using wifi. *SIGCOMM Comput. Commun. Rev. 45*, 4 (Aug. 2015), 269–282.

13. Lazik, P., Rajagopal, N., Shih, O., Sinopoli, B., and Rowe, A. Alps: A bluetooth and ultrasound platform for mapping and localization. In *Proc. SenSys '15*, ACM (2015), 73–84.

14. Lazik, P., and Rowe, A. Indoor pseudo-ranging of mobile devices using ultrasonic chirps. In *Proc. SenSys '12*, ACM (2012), 99–112.

15. Li, M., and Kobbelt, L. Dynamic tiling display: Building an interactive display surface using multiple mobile devices. In *Proc. MUM '12*, ACM (2012), 24:1–24:4.

16. Liu, H., Gan, Y., Yang, J., Sidhom, S., Wang, Y., Chen, Y., and Ye, F. Push the limit of wifi based localization for smartphones. In *Proc. Mobicom '12*, ACM (2012), 305–316.

17. Liu, K., Liu, X., and Li, X. Guoguo: Enabling fine-grained indoor localization via smartphone. In *Proc. MobiSys '13*, ACM (2013), 235–248.

18. Mandal, A., Lopes, C. V., Givargis, T., Haghighat, A., R. Jurdak aMandal, A., Lopes, C. V., Givargis, T., Haghighat, A., Jurdak, R., and Baldi, P. Beep: 3d indoor positioning using audible sound. In *Proc. CCNC '05*, IEEE (2005), 348–353.

19. Medina, C., Segura, J. C., and De la Torre, A. Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy. *Sensors 13*, 3 (2013), 3501–3526.

20. Otsason, V., Varshavsky, A., LaMarca, A., and de Lara, E. Accurate gsm indoor localization. In *Proc. UbiComp '05*, Springer (2005), 141–158.

21. Peng, C., Shen, G., Zhang, Y., Li, Y., and Tan, K. Beepbeep: A high accuracy acoustic ranging system using cots mobile devices. In *Proc. SenSys '07*, ACM (2007), 1–14.

22. Priyantha, N. B., Chakraborty, A., and Balakrishnan, H. The cricket location-support system. In *Proc. MobiCom '00*, ACM (2000), 32–43.

23. Qiu, J., Chu, D., Meng, X., and Moscibroda, T. On the feasibility of real-time phone-to-phone 3d localization. In *Proc. SenSys '11*, SenSys '11, ACM (2011), 190–203.

24. Rajagopal, N., Lazik, P., and Rowe, A. Visual light landmarks for mobile devices. In *Proc. IPSN '14*, IEEE (2014), 249–260.

25. Rishabh, I., Kimber, D., and Adcock, J. Indoor localization using controlled ambient sounds. In *Proc. IPIN '12*, IEEE (2012), 1–10.

26. Rädle, R., Jetter, H.-C., Marquardt, N., Reiterer, H., and Rogers, Y. HuddleLamp: Spatially-aware mobile displays for ad-hoc around-the-table collaboration. In *Proc. ITS '14*, ACM (2014), 45–54.

27. Spindler, M., Büschel, W., Winkler, C., and Dachselt, R. Tangible displays for the masses: Spatial interaction with handheld displays by using consumer depth cameras. *Personal Ubiquitous Comput. 18*, 5 (June 2014), 1213–1225.

28. Tarzia, S. P., Dinda, P. A., Dick, R. P., and Memik, G. Indoor localization without infrastructure using the acoustic background spectrum. In *Proc. MobiSys '11*, ACM (2011), 155–168.

29. Tung, Y.-C., and Shin, K. G. Echotag: accurate infrastructure-free indoor location tagging with smartphones. In *Proc. MobiCom '15*, ACM (2015), 525–536.

30. Ullmer, B., and Ishii, H. The metadesk: Models and prototypes for tangible user interfaces. In *Proc. UIST '97*, ACM (1997), 223–232.

31. von Zadow, U., Büschel, W., Langner, R., and Dachselt, R. Sleed: Using a sleeve display to interact with touch-sensitive display walls. In *Proc. ITS '14*, ACM (2014), 129–138.

32. Woźniak, P., Lischke, L., Schmidt, B., Zhao, S., and Fjeld, M. Thaddeus: A dual device interaction space for exploring information visualisation. In *Proc. NordiCHI '14*, ACM (2014), 41–50.

33. Xiong, J., and Jamieson, K. Towards fine-grained radio-based indoor location. In *Proc. HotMobile '12*, ACM (2012), 13:1–13:6.

34. Yamasaki, R., Ogino, A., Tamaki, T., Uta, T., Matsuzawa, N., and Kato, T. Tdoa location system for ieee 802.11 b wlan. In *Proc. WCNC '05*, vol. 4, IEEE (2005), 2338–2343.

35. Youssef, M., and Agrawala, A. The horus wlan location determination system. In *Proc. MobiSys '05*, ACM (2005), 205–218.

A Proposal for a Taxonomy to Survey the State-of-the-Art of Multi-Device Technologies

Roman Rädle
Digital Design and Information Studies
Aarhus University
roman.raedle@cc.au.dk

Clemens N. Klokmose
Digital Design and Information Studies
Aarhus University
clemens@cavi.au.dk

Hans-Christian Jetter
Human-Centered Computing
University of Applied Sciences Upper Austria
hans-christian.jetter@fh-hagenberg.at

Abstract
Keeping track and gaining an overview of multi-device technologies is an increasingly complex task and judging their quality is challenging. For example, it requires additional effort to gather often scattered information from various sources. This position paper proposes a taxonomy to survey state-of-the-art of multi-device technologies. It contributes nine dimensions that allow to categorize and compare technologies by their individual strengths and weaknesses. It further seeks to start an initiative that calls researchers and practitioners to contribute information on multi-device technologies to an open access catalog. This catalog ought to be the de facto standard and first source to consult when designing, implementing, and studying multi-device interactions.

Author Keywords
taxonomy; dimensions; multi-device; cross-device; multi-surface; cross-surface; enabling technology; open access catalog; comparison, discussion.

ACM Classification Keywords
H.5.2 [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction & Motivation

Research and industry share the growing interest in multi-{device,surface}[1] and in particular in cross-{device,surface} interactions. For example, interests include (i) development of technologies enabling multi-device interactions (e.g., software APIs [27, 5, 13, 23], hardware technologies [17], and hybrids between software and hardware [19, 10]), (ii) design of interaction techniques for multi-device ecologies [4], and (iii) user studies showing benefits when using multiple devices sequentially or in parallel (e.g., active reading [3], sense-making [9, 20, 26], collaborative data analysis [7], collaborative shared workspaces [25], content curation of historical artefacts [1], and audience engagement in presentations [2]). At the same time, industry has started integrating multi-device functionality into the core of operating systems (e.g., Apple Handoff or Microsoft Continuum). New hardware appear at a steady rate that detects nearby devices and services (i.e. Estimote Beacons) together with software APIs to enable novel multi-device experiences (i.e. Google Nearby[2] or Mozilla FlyWeb[3]).

However, the knowledge on multi-device technologies is scattered and stored across different sources such as online databases, platforms, and repositories. For example, digital libraries such as the ACM Digital Library hold a vast amount of research papers, journals, or articles. Other sources are patent databases, online blogs, newsfeeds, and source code repositories such as Github or Bitbucket. Moreover, it is challenging to choose appropriate technology for a multi-device system design because of missing guidelines and metrics defining technology tradeoffs. In this position paper, we propose a taxonomy to survey state-of-the-art of multi-device technologies and evaluate their tradeoffs. We further call researchers and practitioners likewise to collect information on multi-device technologies and contribute this information to an open access catalog[4]. This catalog is supposed to become the de facto standard and first source to consult when designing, implementing, and studying multi-device experiences. It will help the HCI community to understand multi-device technology tradeoffs better and will open up possibilities for them to discuss these technologies alongside their benefits and shortcomings.

[1] The terms *-device and *-surface are often used synonymously.

[2] Google Nearby Project Website – https://developers.google.com/nearby/ (last accessed: September, 29, 2016

[3] Mozilla FlyWeb Project Website – https://flyweb.github.io/ (last accessed: September 29, 2016

Goal and Expected Contributions

Despite existing studies on multi-device use [22, 12] and surveys on multi-device challenges [8] or device association techniques [6], literature is lacking appropriate tools to assess the quality of multi-device technologies. It is, therefore, challenging for designers to select appropriate technology when designing multi-device systems. For example, the design of a system might afford the use of a particular technology because of the requirement for spatial tracking of devices. However, benefits of technologies also often come at a cost such as additional instrumentation of an environment.

We want to support this decision making by providing a taxonomy for assessing multi-device technologies and report results in a publicly accessible catalog. This will inform researchers and practitioners about the state-of-the-art of multi-device technologies and allow them to understand technology tradeoffs. It will further overview and detail on multi-device application domains, interaction techniques designed for multi-device & cross-device use, multi-device

[4] The open access catalog is inspired by the iLab Cookbook – http://grouplab.cpsc.ucalgary.ca/cookbook/ (last accessed: September 29, 2016).

user studies, and methods to systematically evaluate the past, present, and future technologies. Ideally, such an open access catalog will also guide future research by revealing remaining research challenges and research gaps.

In contrast to Grubert et al. [8], we narrow our focus to technology and only Weiserian devices [24]. This includes the three broad device types: *tabs* (“inch-scale”), *pads* (“foot-scale”), and *boards* (“yard-scale”). It also includes both traditional personal devices, but also shared devices such as wall-displays or interactive whiteboards. We opted for this focus because we believe that a too broad view on technology will impede the development of a taxonomy and eventually not allow for proper technology comparison. However, related technologies such as virtual reality (VR), augmented reality (AR), and mixed reality (MR) could potentially be added at a later stage when possible. It is important to notice that this position paper makes neither claim to feature a final set of taxonomy dimensions nor does it exhibit a complete summary of the state-of-the-art multi-device research. It is a starting point, which solicits the multi-device community to join forces and calls for their contribution.

In the remainder of this position paper, we present our initial set of taxonomy dimensions to survey multi-device technologies. Ideally, these dimensions offer a structured way to assess technologies and help to identify their commonalities, differences, and tradeoffs. Eventually, this structured analysis will reveal trends, challenges, and gaps in multi-device development, inform research questions and thrive future studies in this area.

Taxonomy and Dimensions

To categorize and rate enabling technologies, we created an initial set of nine dimensions for assessing different qualities: *device setup*, *environment instrumentation*, *tracking fidelity*, *tracking space & dimension*, *tracking quality*, *tracking versatility*, *tracking technology*, *hardware & material cost*, and *availability*. Each dimension is briefly described in the following sections (see Table 1 for an overview); dimensions are presented in no particular order. Also, we would like to take the opportunity to present and discuss the taxonomy and its nine dimensions at the cross-surface workshop. As a result of the workshop, a final taxonomy and its dimensions alongside example technologies will be submitted as an article to the special issue on “Interaction with Device Ecologies in the Wild”.

Device Setup

An obvious aspect of multi-device technologies is device setup. An ideal technology works without prior configuration of the target hardware (*zero-conf*). It should run on commodity and off-the-shelf hardware [19]. For example, multi-device systems [5, 13, 23, 18] implemented on modern web standards operate in a browser and do not require the user to install additional software (*zero-install*).

Environment Instrumentation

To track spatial locations and orientation of devices, rooms or environments are often equipped with additional tracking hardware. For example, the Proximity Toolkit [15] requires a *fixed* instrumentation of an entire space with motion capture cameras such as Vicon or OptiTrack. Other systems like HuddleLamp [19] also require to instrument the environment. However, their setup is portable and therefore considered *semi-fixed*. Software frameworks like Weave [5] or Connichiwa [23] allow for development of *mobile* multi-device experiences.

Tracking Fidelity

The tracking fidelity of devices often resonates with aforementioned instrumentation of the environment. The Prox-

Tracking Quality from HuddleLamp [19]

Precision is the standard deviation of the tracked position and orientation of a fixed tablet over time and thus measures noise and jittering (units: mm or degree).

Accuracy is the spatial accuracy of a tablet's position compared to a ground truth (unit: mm).

Reliability is the percentage of frames in which a present tablet was tracked (its tracking state is 'true') during measurement (unit: %).

Dimension	Values
Device Setup	zero-conf, zero-install
Environment Instrumentation	fixed, semi-fixed, mobile
Tracking Fidelity	spatially-agnostic, synchronous gestures, spatially-aware
Tracking Space & Dimension	2D, 3D and small-scale, mid-scale, large-scale
Tracking Quality	accuracy, precision, reliability, and close, medium, far
Tracking Versatility	user identity, body, limbs, head, hand
Tracking Technology	computer vision, Bluetooth LE, infrared, (in-)audible sounds, radio signals, WLAN
Hardware & Material Cost	consumer hardware, special hardware, research prototype
Availability	open source, closed source, proprietary

Table 1: Taxonomy including nine dimensions to assess multi-device technologies.

imity Toolkit [15] and HuddleLamp [19] represent two examples of technologies that enable development of *spatially-aware* cross-device interactions. Connichiwa [23] implements *synchronous gestures* for display stitching and others like Webstrates [13] and XDBrowser [18] are yet *spatially-agnostic*. The survey paper by Chong et al. on spontaneous device associations [6] could inspire further values for this dimension.

Tracking Space & Dimension
Technologies that allows for spatial tracking can be further divided into tracking space and tracking dimension. While HuddleLamp [19] is restricted to a *small-scale* space approx. 100$\times$60cm and only tracks devices in *2D* space, Tracko [11] (*mid-scale*) and the Proximity Toolkit [15] (*large-scale*) allow spatial tracking in *3D* space and to a much larger extent. Latter depends on the motion capture camera setup, which often spans an interaction space of several cubic meters.

Tracking Quality
User experience of multi-device and cross-device interaction heavily relies on the tracking quality. For example, HuddleLamp [19] introduces the three measures *accuracy*, *precision*, and *reliability* and uses them to assess the quality of their hybrid sensing, which exploits optical characteristics and computer vision to track mobile devices (see sidebar for definition of HuddleLamp's quality measures). Tracko [11], as another example, shows how to categorize tracking quality based on three discrete distances: *close* ($<$0.5m), *medium* (0.5m-1.0m), and *far* ($>$1.0m).

Tracking Versatility
In addition to device tracking, some technologies allow tracking of *users*, their *bodies* or *limbs*. HuddleLamp, for example, tracks users' *hands* for mid-air cross-device interactions [19]. The GroupTogether [16] system tracks users' *body* & *head* position and orientation to guide user interaction using F-formations theory.

Tracking Technology
Most tracking techniques exploit internal device sensors and/or external hardware to track devices. HuddleLamp uses an RGB-D camera and *computer vision* (CV) for its hybrid sensing tracking [19]. Sifteo cubes use *infrared light* (IR) to detect neighboring cubes [17], Connichiwa exploits

Bluetooth Low Energy (BT LE) signals to calculate distance to other devices [23], Tracko uses *in-audible sounds* and *microphones* for triangulation and location detection of devices [11], GroupTogether uses short-range communciation technology and *radio signal* trilateration [16], and ProxiMagic uses *WLAN* for proximity detection (proximity-adaptive HTTP responses) [14].

Hardware & Material Cost
For most controlled lab research, the cost for hardware & material necessary for multi-device interaction is not a primary concern. From our experience and working with public and academic libraries [21] and schools [20], the cost of technology, however, is often an important factor for "in-the-wild" deployments (e.g., to reduce operating costs or costs for replacement in case of vandalism). An ideal system exploits users' devices and encourages them to "bring-your-own-device" (BYOD). It, therefore, should work without or only a little instrumentation of the environment. Such concerns might resonate with "in-the-wild" deployments at other public spaces such as community centers and museums.

Availability
The use of multi-device technology is often limited by the availability of source code. The Tracko paper [11], for example, describes the tracking algorithms, but the source code is *propriety* and additional effort is needed to re-implement the tracking. Likewise, research projects keep source code offline (*closed source*). Recent projects, however, published code as *open source* such as Proximity Toolkit [15], WatchConnect [10], Webstrates [13], and HuddleLamp [19].

Summary
We proposed a taxonomy to survey state-of-the-art of multi-device technologies and to assess their quality using the nine dimensions: *device setup*, *environment instrumentation*, *tracking fidelity*, *tracking space & dimension*, *tracking quality*, *tracking versatility*, *tracking technology*, *hardware & material cost*, and *availability*. This taxonomy allows to categorize and compare multi-device technologies, help to identify their commonalities, differences, and tradeoffs. We further call researchers and practitioners to contribute information on multi-device technologies to an open access catalog. This catalog can become the first source to consult when designing, implementing, and studying multi-device interactions.

REFERENCES

1. Fredrik Brudy, Steven Houben, Nicolai Marquardt, and Yvonne Rogers. 2016. CurationSpace: Cross-Device Content Curation Using Instrumental Interaction. In *To appear: In Proceedings of the 11th ACM International Conference on Interactive Surfaces and Spaces - ISS '16*. New York, NY, USA.

2. Debaleena Chattopadhyay, Kenton O'Hara, Sean Rintel, and Roman Rädle. 2016. Office Social: Presentation Interactivity for Nearby Devices. In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, NY, USA.

3. Nicholas Chen, François Guimbretière, and Abigail Sellen. 2013. Graduate student use of a multi-slate reading system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, New York, New York, USA, 1799.

4. Xiang 'Anthony' Chen, Tovi Grossman, Daniel J. Wigdor, and George Fitzmaurice. 2014. Duet: exploring joint interactions on a smart phone and a smart watch. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 159–168.

5. Pei-yu Peggy Chi and Yang Li. 2015. Weave: Scripting Cross-Device Wearable Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM, New York, NY, USA, 3923–3932.

6. Ming Ki Chong, Rene Mayrhofer, and Hans Gellersen. 2014. A Survey of User Interaction for Spontaneous Device Association. *Comput. Surveys* 47, 1 (may 2014), 1–40.

7. Johannes Fuchs, Roman Rädle, Dominik Sacha, Fabian Fischer, and Andreas Stoffel. 2014. Collaborative data analysis with smart tangible devices. In *IS&T/SPIE Electronic Imaging*, Pak Chung Wong, David L. Kao, Ming C. Hao, and Chaomei Chen (Eds.). International Society for Optics and Photonics.

8. Jens Grubert, Matthias Kranz, and Aaron Quigley. 2016. Challenges in Mobile Multi-Device Ecosystems. (may 2016).

9. Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: enabling and understanding cross-device interaction. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14 (CHI '14)*. ACM Press, New York, New York, USA, 2773–2782.

10. Steven Houben and Nicolai Marquardt. 2015. WatchConnect: A Toolkit for Prototyping Smartwatch-Centric Cross-Device Applications. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 1247–1256.

11. Haojian Jin, Christian Holz, and Kasper Hornbæk. 2015. Tracko: Ad-hoc Mobile 3D Tracking Using Bluetooth Low Energy and Inaudible Signals for Cross-Device Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. ACM Press, New York, New York, USA, 147–156.

12. Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 3903–3912.

13. Clemens N. Klokmose, James R. Eagan, Siemen Baader, Wendy Mackay, and Michel Beaudouin-Lafon. 2015. Webstrates: Shareable Dynamic Media. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. ACM Press, New York, New York, USA, 280–290.

14. Clemens Nylandsted Klokmose, Matthias Korn, and Henrik Blunck. 2014. WiFi proximity detection in mobile web applications. In *Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems - EICS '14*. ACM Press, New York, New York, USA, 123–128.

15. Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. 2011. The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. ACM Press, New York, New York, USA, 315.

16. Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, New York, New York, USA, 13.

17. David Merrill, Jeevan Kalanithi, and Pattie Maes. 2007. Siftables: towards sensor network user interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction - TEI '07 (TEI '07)*. ACM Press, New York, New York, USA, 75.

18. Michael Nebeling and Anind K. Dey. 2016. XDBrowser: User-Defined Cross-Device Web Page Designs. In *To Appear: Proceedings of the 34th annual ACM conference on Human factors in computing systems - CHI '16*. ACM Press, New York, NY, USA.

19. Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS '14*. ACM Press, New York, New York, USA, 45–54.

20. Roman Rädle, Hans-Christian Jetter, and Harald Reiterer. 2013. TwisterSearch: A distributed user interface for collaborative Web search. In *Distributed User Interfaces: Collaboration and Usability*, María D. Lozano, Jose A. Gallud, Ricardo Tesoriero, and Victor M.R. Penichet (Eds.). Springer London, London, UK, Chapter 5, 4–7.

21. Harald Reiterer, Roman Rädle, Simon Butscher, and Jens Müller. 2016. Blended Library – neue Zugangswege zu den Inhalten wissenschaftlicher und öffentlicher Bibliotheken. *Bibliothek Forschung und Praxis* 40, 1 (jan 2016).

22. Stephanie Santosa and Daniel Wigdor. 2013. A field study of multi-device workflows in distributed workspaces. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*. ACM Press, New York, New York, USA, 63.

23. Mario Schreiner, Roman Rädle, Hans-Christian Jetter, and Harald Reiterer. 2015. Connichiwa: A Framework for Cross-Device Web Applications. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. ACM Press, New York, New York, USA, 2163–2168.

24. Mark Weiser. 1991. The computer for the 21st century. *Scientific American* 265, 3 (sep 1991), 94–104.

25. Daniel Wigdor, Hao Jiang, Clifton Forlines, Michelle Borkin, and Chia Shen. 2009. WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09.* ACM Press, New York, New York, USA, 1237.

26. Johannes Zagermann, Ulrike Pfeil, Roman Rädle, Hans-Christian Jetter, Clemens Klokmose, and Harald Reiterer. 2016. When Tablets meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 5470–5481.

27. Michael Zöllner, Hans-Christian Jetter, and Harald Reiterer. 2011. ZOIL: A Design Paradigm and Software Framework for Post-WIMP Distributed User Interfaces. In *Distributed User Interfaces: Designing Interfaces for the Distributed Ecosystem*, José A. Gallud, Ricardo Tesoriero, and Victor M.R. Penichet (Eds.). Springer London, London, 87–94.

Interactions in a Human-Scale Immersive Environment: the *CRAIVE*-Lab

Gyanendra Sharma
Department of Computer Science
Rensselaer Polytechnic Institute
sharmg3@rpi.edu

Jonas Braasch
School of Architecture
Rensselaer Polytechnic Institute
braasj@rpi.edu

Richard J. Radke
Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute
rjradke@ecse.rpi.edu

Abstract
We describe interfaces and visualizations in the CRAIVE (Collaborative Research Augmented Immersive Virtual Environment) Lab, an interactive human scale immersive environment at Rensselaer Polytechnic Institute. We describe the physical infrastructure and software architecture of the CRAIVE-Lab, and present two immersive scenarios within it. The first is "person following", which allows a person walking inside the immersive space to be tracked by simple objects on the screen. This was implemented as a proof of concept of the overall system, which includes visual tracking from an overhead array of cameras, communication of the tracking results, and large-scale projection and visualization. The second "smart presentation" scenario features multimedia on the screen that reacts to the position of a person walking around the environment by playing or pausing automatically, and additionally supports real-time speech-to-text transcription. Our goal is to continue research in natural human interactions in this large environment, without requiring user-worn devices for tracking or speech recording.

Author Keywords
Interaction; Immersive Environment; Person Tracking; Multimedia

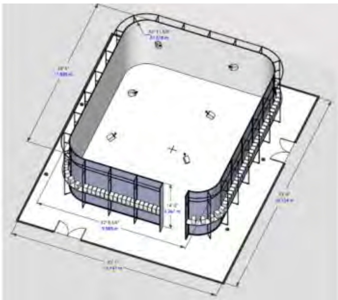

Figure 1. The physical and infrastructural specifications of the CRAIVE-Lab. The lab is equipped with 8 HD video projectors and 128 loudspeakers behind a 5m tall, acoustically transparent micro-perforated screen. 6 downward-pointed cameras are mounted in the ceiling of the space for continuous tracking visual tracking of participants. Figure courtesy J. Parkman Carter.

Figure 2. Snapshot of the CRAIVE lab with a panoramic image on display.

ACM Classification Keywords
H.5.m. Information interfaces and presentation

Introduction
Large-scale display technologies have several interesting applications including the presentation of high-resolution imagery for geospatial, scientific visualization, and command-and-control scenarios [7]. When large displays enclose a space to create an immersive environment, it is not viable for users to interact with the displays using traditional interfaces like a mouse or keyboard. The scale of the environment makes it difficult to carry out even basic tasks such as finding the cursor or clicking on a certain part of the screen. In addition, even if visualizations, imagery or multimedia can be efficiently presented on the large screen, exploring them without using invasive technologies such as wearable sensors can be difficult. In this paper, we present the CRAIVE (Collaborative Research Augmented Immersive Virtual Environment) Lab, a human-scale, occupant-aware immersive environment at Rensselaer Polytechnic Institute. Our current work focuses on creating interactive interfaces in the CRAIVE-Lab by automatically detecting the spatial and temporal positions of the participants. Ultimately, this tracking will enable personalized communication between the smart environment and each user through mobile devices and handheld surfaces.

The CRAIVE-Lab, illustrated in Figures 1 and 2, has a 360-degree front-projected screen equipped with 8 1200x1920 resolution projectors along with a network of 6 overhead 1280x960 resolution cameras used for visual tracking of participants. The floor space enclosed within the screen is rectangular with curved corners and has a length of approximately 12 meters and width of 10 meters. The screen is approximately 5 meters in height, and has an effective resolution of 1200x14500 pixels. The projectors are positioned to avoid casting user shadows onto the screen unless they are very near to it.

Here, we describe several user interfaces to the CRAIVE-Lab that require neither wearable sensors nor direct interaction with the system through a mouse or keyboard. We use computer vision algorithms for user tracking along with the Websocket message passing interface to realize an occupant-aware environment. For example, in the "smart presentation" scenario, a user simply has to stand close to a multimedia clip in the physical space, activating the clip by their presence and automatically transcribing their speech in the appropriate location.

The remainder of the paper discusses the integral building blocks of the CRAIVE-Lab system, including visual tracking of participants from above, screen and floor coordinate system calibration, and system architecture, resulting in real-time systems for "person following" and "smart presentation" scenarios.

Related Work

Human scale immersive interactive environments have recently become a research focus in industry, government, and academia. Natural interactions in the context of these environments is a critical issue. Our approach focuses on natural and non-invasive techniques for building occupant-aware spaces. Zabulis et al. [11] explored the problem in a similar way, supporting natural interaction in a large screen display environment.

Bradel et al. [3] studied how multiple users can collaborate and interact using large high resolution displays, though they were not in the context of an immersive environment. However, it is useful to understand how large displays affect collaboration, which is an important issue for our system. North and North [8] studied user experiences based on virtual reality and immersive environments, investigating factors that affect the sense of presence and immersion. One of the major contributions of this study is the conclusion that immersive systems provide users with a heightened sense of presence. This was one of the motivations for the work presented in this paper: to create interfaces that aid in creating meaningful immersive experiences.

Ardito et al. [1] presented a survey related to interactions with large displays. While the testbed for the survey involved simple large screen displays instead of extremely large walls as in our system, several of the surveyed interaction techniques could be used for future interfaces in the CRAIVE-Lab. Interaction techniques involving handheld devices, e.g., [4, 5, 6, 10], are also well within the scope of our future work.

System Architecture

Figure 3 illustrates the software architecture of the CRAIVE-Lab system. The 6 overhead cameras provide a live feed to the tracking algorithm, which returns the global (x, y) floor coordinate for each user inside the CRAIVE-Lab. Using a Websocket client, the coordinates are sent to a Nodejs server hosted in the same node (i.e., computer). This server in turn communicates via Websocket to the HTML client hosted in the projector node. User-worn lapel microphones communicate

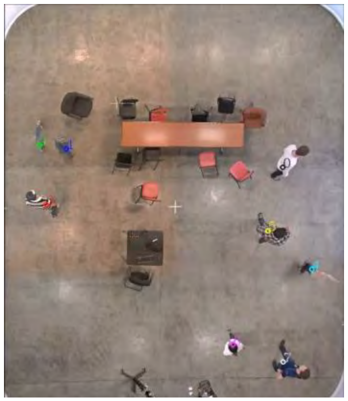

Figure 4. A snapshot of the live visual tracking system in the CRAIVE-Lab. A unique ID (colored number) is assigned to each occupant, which allows the system to keep track of all participants in the environment.

directly with the projector node, using a web client to transcribe speech to text using the Google speech API.

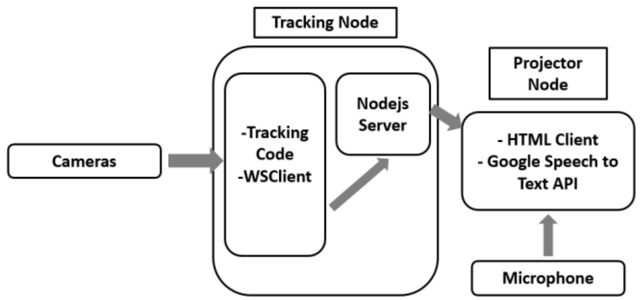

Figure 3. Block diagram of the overall system architecture.

Tracking

The CRAIVE-Lab visual tracking algorithm operates in real-time on each of the 6 cameras, fusing the results into target positions in a global coordinate frame, as illustrated in Figure 4. The six cameras provide full coverage of the CRAIVE-Lab floor, allowing for continuous tracking across any region in the room.

Users are detected inside the space using a background subtraction method. A multi-threaded approach is used to track people in individual cameras. At each moment, the individual camera feeds are combined into a mosaic covering the entire space with proper alignment of the overlapping regions between neighboring cameras. This alignment involves camera-to-camera homographies relating the ground planes in each viewpoint, calibrated at system setup using feature correspondences in the overlap regions between fields of view. In this way, the local tracking coordinates returned by each camera are mapped into global coordinates for the floor.

Calibration

To use spatial and temporal motion of a user in the space to drive the interactive visualization on the screen, a mapping between floor and screen coordinates is required. The mapping should enable visual cues on the screen to aid the user in intuitively translating where he/she is on the floor to a location on the screen.

The first issue is determining a unified screen coordinate system, taking into account that the display is generated from 8 overlapping projected images. We use the PixelWarp software package [5] to handle the 8-projector display as a virtual 14500-pixel-wide single image. The calibration process involves both geometric warps and color blends in the overlap regions, so the projection is virtually seamless to participants. The projector-to-projector mappings were created using a meticulous manual calibration process when the system was commissioned.

We determined the mapping from floor coordinates to screen coordinates by positioning a user at 200-pixel (roughly 0.7 m) intervals at a fixed distance from the display wall along its entire 14500-pixel-wide extent, and recording the corresponding (x, y) locations in floor coordinates. This mapping is generally linear, but must take into account the curved regions at the corners of the projection screen. Ultimately, the user's floor x coordinate drives the mapping to Screens 1 and 3, and their y coordinate drives the mapping to Screens 2 and 4, as illustrated in Figures 5 and 6.

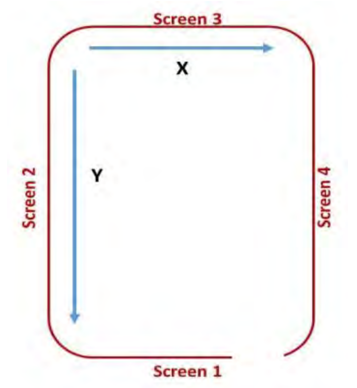

Figure 6. Illustration of the CRAIVE-Lab screen and floor space. The origin of the floor coordinate system is at the upper left corner and coordinates increase as indicated by the arrows. Note that the screen numbers simply indicate locations of content for the purposes of this paper; the screen itself is a single long display surface.

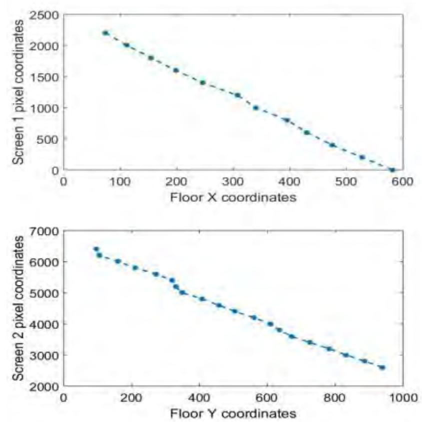

Figure 5. Floor coordinates vs screen pixel coordinates for Screens 1 and 2. Screens 3 and 4 follow similar patterns.

Person Following

As shown in Figure 7, we aim to provide a visual cue to a user walking inside the immersive space. Our decision about the visualization artifact was made with the goal of providing unambiguous feedback about the user's spatial location. For this purpose, we took a "crosshair"-like approach for visualization. A user walking around the space will be followed by four filled circles, one on each screen, such that connecting opposite circles pinpoints the user in the interior of the room. This is illustrated from overhead in Figure 8.

Hence, all four sides of the environment provide feedback to the user regarding his/her location in the space. However, in order to avoid ambiguity as well as to enhance occupant awareness, the circles are resized dynamically. That is, we maintain an inversely proportional relationship between the user's distance to

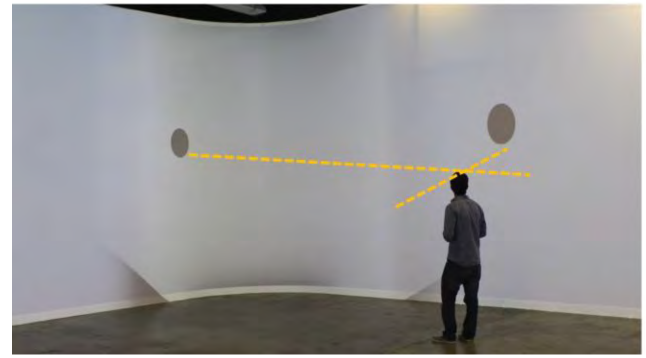

Figure 7. Snapshot of "person following" in the CRAIVE-Lab, with lines added for illustration. Circles with dynamic sizes and locations across the screen are used as visual cues for a user walking around the space and act as "crosshairs". A video link is available at https://youtu.be/o6cF7_Gyz3Y.

a screen and the circle size on that screen (i.e., the circle on the screen closest to the user is the largest). Also, in order to reduce noise from the tracking algorithm, a median filter is applied across 20 consecutive frames so that the circles on the screens do not have too much fluctuation.

Smart Presentation / Multimedia

We next present a more advanced occupant-aware scenario that uses the same underlying tracking algorithm. This design extends the "person following" idea to show how the existing technology can be applied for intuitive, interactive presentations. Since we know the tracking coordinates and their corresponding mapping to screen coordinates as discussed in the previous sections, we designed a multimedia presentation system controlled by the spatial location of the user. The idea is to have several large videos at different positions along the screen. Each video plays

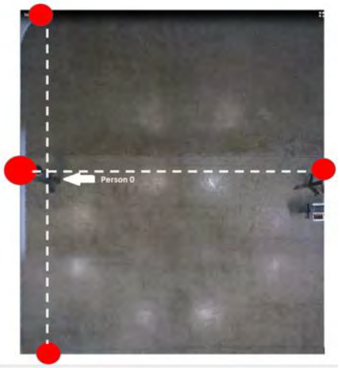

Figure 8. Four circles pinpoint the user's location in the room interior. Here, a mosaiced overhead view of the space is shown with a person in the left side. Four circles on each of the sides of the screen are automatically drawn, such that the virtual dotted lines intersect at the user's position. The screen to which the person is closest has the largest circle.

only when the user is right in front of the video and sufficiently close to the screen. In addition, the user wears a lapel microphone, allowing him/her to talk about the presentation/multimedia; this speech is automatically transcribed and displayed above the corresponding video. Any comments made by the user while outside the mapped floor region of the multimedia content is ignored. This allows the user to move around the space while controlling the content on the screen without making any conventional active interactions such as mouse clicks. An illustration of this idea is shown in Figure 9.

Figure 9. Snapshot of the "smart presentation". Here, the user has moved to the first video on the left. All other videos are paused as the user is not in the mapped floor region of either of the two videos, while the first one is playing along with the audio transcription appearing above it. A video link is available at: https://youtu.be/VbXMB97HwvM

A prototype of this system was implemented in a separate space similar to the CRAIVE-Lab during the IBM Cognitive Colloquium in November 2015. Various research projects were presented using this system, in which the presenter walked around the room explaining the projects while the videos played and paused automatically along with automatic speech transcription. The response from the attending participants was very positive and successfully showcased the occupant-awareness of the room and its ability to support natural user interactions.

Conclusions and Future Work

In near-term future work, we plan to integrate various cross-surface devices such as tablets with the large screen. A user could easily use his/her mobile surface to share local content onto the display wall; similar to the "smart presentation" interface, the content's screen location could be synchronized with the user's spatial location. A reverse scenario of content sharing from the display wall to specific user devices could also be realized using the system we have in place.

The CRAIVE-Lab can be further improved to include various gestural or voice interactions apart from the spatially controlled interactions. For example, floor-mounted cameras and range sensors like the Microsoft Kinect can capture users' gestures, head poses, and facial expressions for finer control of presentation elements. In addition, multimedia elements in the environment could be controlled directly from speech understanding, preferably using ambient microphones instead of lapel-worn wireless microphones.

Our ultimate goal and current work in the CRAIVE-Lab focuses on making effective use of the large space with respect to service systems that aid both individual and group interactions, interfacing with advanced cognitive computing algorithms "behind the screen". Ultimately, we envision an intelligent environment that understands its occupants' locations, movement, speech, vocabulary, and intentions, and is able to actively facilitate meetings, aware of the day's agenda and action items, the participants and their roles, and what happened previously. The environment could keep meetings on track, assess participation and progress towards goals, maintain action items, summarize discussion, and mediate brainstorming.

Acknowledgements

The construction of the CRAIVE-Lab was supported by the National Science Foundation and by Rensselaer Polytechnic Institute [2]. Robert Carney and J. Parkman Carter worked with the CRAIVE-Lab project team to plan and direct the installation of the CRAIVE-Lab screen and loudspeakers and to calibrate the system. The metal frame for the CRAIVE-lab was built by Adirondack Studios. Jacob Braun installed the cameras and implemented the visual tracking algorithm based on an original algorithm by Ziyan Wu. The research in this paper was partially supported by NSF's PFI:BIC program under award IIP-1631674, and by IBM Research via Rensselaer's Cognitive and Immersive Systems Laboratory.

References

[1] Ardito, C., Buono, P., Francesca, C. and Desolada, G., Interaction with large displays: A survey. *ACM Computing Survey*, 47, 2015.

[2] Braasch, J. (PI), Chang, B. (Co-PI), Cutler, B. (Co-PI), Goebel, J. (Co-PI) and Radke, R. (Co-PI), MRI/Dev.: Collaborative-Research Augmented Immersive Virtual Environment Laboratory (CRAIVE-Lab), NSF/Major Research Instrumentation, 10/12–09/15, CNS-1229391.

[3] Bradel, L., Endert, A., Koch, K., Andrews, C. and North, M., Large high resolution displays for co-located collaborative sensemaking: Display usage and territoriality. *International Journal of Human Computer Studies*, 71:1078-1088, 2013.

[4] Jeon, S., Hwang, J., Kim, G.J. and Billinghurst, M., Interaction techniques in large display environments using handheld devices. *Proceedings – ACM Symposium on Virtual Reality Software and Technology – VRST '06*, 2006.

[5] Malik, S., Ranjan, A., and Balakrishnan, R., Interacting with large displays from a distance with vision-tracked multi-finger gestural input. *Proceedings – ACM symposium on User Interfave Software and Technology – UIST '05*, 2005.

[6] Nancel, M., Chapuis, O., Pietriga, E., Yand, X.D., Pourand, I. and Beaudouin-Lafon, M., High-precision pointing on large wall displays using small handheld devices. *Proceedings – SIGCHI Conference on Human Factors in Computing Systems – CHI '13*, 831-840, 2013.

[7] Ni, T., Schmidt, G.S., Staadt, O.G., Livingston, M.A., Ball, R. and May, R., A survey of large high-resolution display technologies, techniques, and applications. *Proceedings - IEEE Virtual Reality*, page 31, 2006.

[8] North, M.M. and North, S., A comparative study of sense of presence of traditional virtual reality and immersive environments. *Australasian Journal of Information Systems*, 20:1-15, 2016.

[9] Pixelwarp warp and blend software. http://pixelwix.com/6-warp-and-blend-software

[10] Vogel, D., and Balakrishnan, R., Distant freehand pointing and clicking on very large, high resolution displays. *Proceedings of the ACM symposium on User Interface Software and Technology – UIST '05*, 2005.

[11] Zabulis, X., Grammenos, D., Sarmis, T., Tzevanidis, K., Padeleris, P., Koutlemanis, P. and Argyros, A.A., Multicamera human detection and tracking supporting natural interaction with large-scale displays. *Machine Vision and Applications*, 24:319-336, 2013.

A Tool for Multi-Surface Collaborative Sketching

Jorge Luis Pérez-Medina
Université catholique de Louvain
Louvain School of Management
Research Institute
Place des Doyens, 1-B-1348
Louvain-la-Neuve (Belgium)
jorge.perezmedina@uclouvain.be

Jean Vanderdonckt
Université catholique de Louvain
Place des Doyens, 1-B-1348
Louvain-la-Neuve (Belgium)
jean.vanderdockt@uclouvain.be

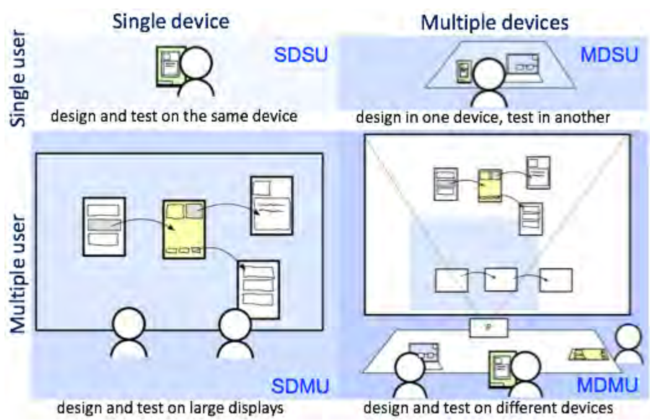

Figure 1: Four physical configurations of a Collaborative User Centered Design Method considering multiple users and devices: SDSU=single device and single user; SDMU = single device and multiple users; MDSU= multiple devices and single user; MDSD= multiple devices and multiple users.

Presented at the Cross-Surface'16 workshop, in conjunction with ACM ISS'16. November 6, Niagara Falls, Canada.

Abstract
This research paper focus on Collaborative Sketching as a tool to facilitate the User's Interface prototyping with multi-surface required by stakeholders and final users in early stages of the software development when developers have a vague idea in mind of what should the service or system being developed or improved will be. The paper also offers a Collaborative User Centered Design method based on sketching.

Author Keywords
Electronic sketching; Multi-surfaces; Collaborative design; Prototyping.

ACM Classification Keywords
H.5 [INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)]: H.5.2 User Interfaces - Graphical user interfaces (GUI) - Prototyping - User-centered design

Introduction
Software modeling tools are often not used during early steps of the software life-cycle for which they would be more helpful [11]; free-form approaches, such as office tools, white boards and papers arranged on a wall, are frequently used instead [14]. We agree with the recommendations expressed by [14]. They suggest the SE community to pay more attention to use the tried-and-tested methods.

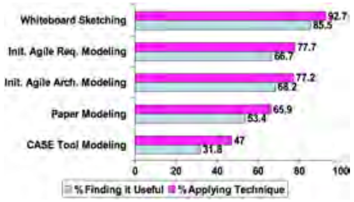

Figure 2: Primary approaches to modeling [1].

A survey realized by [1] summarize the way how effective initial requirements modeling was on agile teams. Figure 2 shows that 92.7% of respondents indicated that those teams were using *Witheboard Sketching* for requirements gathering, and that 85.5% of those teams found the effort worthwhile.

Indeed, these can help to produce more and usable tools and bring together people who understand the activities and needs of developers and other stakeholders to produce User's Interface design throughout the software life-cycle.

The User Interface is the portion of software with which a user directly interacts. Designers can explore the requirements using essential UI prototypes [1]. It represent UI requirements in a technology independent manner, just as essential use case models do for behavioral requirements. Although essential UI prototyping is a great technique, many stakeholders find them too abstract and instead prefer screen sketches and not surprisingly working software [1]. Designers will use the simplest tool which will get the job done. For example, whiteboard diagrams typically suffice for initial requirements models.

We aim at providing support for collaborative and flexible modeling in early steps of a User Centered Design process where stakeholders and final users can collaborate by using multi-surface provided by multi-device. It consists of using sketching in early UI design as an informal and flexible manner for producing any model for requirements and analysis phases. Sketching can also be used by any stakeholders using any device as it is a way of expression which does not require any advanced modelling skills. Sketches are therefore refined to achieve user validation. Iterative UI design tools, used in HCI, generally support electronic sketching, giving more freedom to change sketches and more flexibility in creating and evaluating a design prototype [8]. The UI development process, is usually conducted in HCI. It appears to be eminently informal to support collaborative activities involving stakeholders and final users.

By using sketches designers may be more motivated, more creative, and perhaps better able to address the challenges of creating a more successful design, and thus produce a better design outcome [16]. Prototyping the UI can detect ergonomic issues that will arise even before the first line of code is produced. The prototype is also a valuable communication tool; the operation of a tool can be described to a non-computer user with greater ease. Prototyping is characterized by collaborative interactions that inherently influence how design is actually practiced, therefore a tool to support User Interface Design should ideally be flexible in order to accommodate those design situations, i.e, different activities, people from different areas and using different media and devices. We propose a combination of 4 possible configurations presented in Figure 1 which considers the different users (or roles which contains both designers and users) and devices.

This paper presents a flexible Collaborative User Centered Design (CUCD) method based on sketching to support prototyping along a Software Development process. It promotes flexible communication among stakeholders and end-users in early steps of development when it is vital to understand the system requirements. The remainder of this paper is structured as follows. In section 2, sketching and basic concepts of design by sketching are presented. Section 3 presents some related work. Section 4 describes a (CUCD) method considering that all its actors have potential to participate in the prototyping process. Section 5 introduces GAMBIT, a tool supporting the method. Section 6 presents the results of a pilot experiment. Finally, conclusions and some remarks for future work are presented.

Design by Sketching

Sketches are rapid drawings [3]. Sketching is useful as a design tool. Stakeholders and final users are able to carry out UI design by sketching during collaborative sessions of brainstorming of ideas to understand the requirements.

Design is a hard challenging for designers. Design is essentially a problem of wicked nature[1]; On wicked problem the complexity can be due incomplete or contradictory knowledge, the number of people and opinions involved, the large economic burden, and the interconnected nature of these problems with other problems, i.e. the process of solving it is identical with the process of understanding it.

Scott Ambler in 2003 defines UI Prototyping as an iterative analytical technique in which users are actively involved [1]. With UI prototyping, the final user is able to provide valuable feedback from early stages of development.

In wicked problems, the designer does not have a clear understanding of what to produce. During the initial phases of this process designers have only a vague goal in mind. Sketching stimulates a re-interpretive cycle in the individual designer's idea generation process [13]. The prototyping activity is then complementary to the sketching activity in order to achieve understanding, especially if the artifact being designed has more dimensions than those used by the designer while sketching.

In the HCI domain, UI design shares the same aforementioned properties: the wicked nature, the use of *"cheap"* tools like sketching, the use of low-fidelity models such as prototypes, etc. The main difference is that UI design is included in more general methods such as User Centered Design which has specific roles that attribute different responsibilities to the party involved. The activity of UI sketching has several recognized virtues, mainly for maintaining a flexible representation of a design, which is reckoned to foster creativity [2].

Related Work

In literature review, we found many software development frameworks like [6], interactive applications and academic research supporting sketching activities. Efforts address mainly tools and methods. [12] presents a extensive comparison of sketching tools considering UI design by sketching and Collaborative Design.

FlexiSketch TEAM [15] is a solution for collaborative, model-based sketching that allows multiple users using their own tables to work on the same model sketch simultaneously and use lightweight meta-modeling mechanics to collaboratively define custom notations on the fly. It considers a mechanism for quickly gathering and validating requirements, for which User Centered Design (UCD) can be used.

[10] concludes that paper prototypes have been found to be useful in practice in many more ways than just for usability testing. [9] presents a vision of how UCD and agile practices coexist in a development team. It puts forward the relationship between the agile practices and UI design, for example: on agile practices there is little notion of a "UI designer" specialty, but, on HCI domain, UI designers practices may be more of a *"team effort"*. Agile projects could require UI design, which typically iterates only on the UI using low technology prototypes. Finally, [5] shows the importance of using UCD as an early practice in the software development lifecycle. The inclusion of the collaborative method within UCD and Agile is aligned with previous proposals, such as Communication-Centered design and Extreme Designing [4], where the focus is on the communication between designers and users. Those approaches advocate that instead of having user stories and sketches as the only design documentation, the use of a streamlined structuring representation that binds together the stories and sketches would benefit the process of designing interfaces.

A Collaborative User Centered Design method based on sketching

This section presents a (CUCD) method for prototyping based on sketching, enabling fast, flexible, intuitive and reusable prototype. The method is derived from the observation of UCD practices (like Brainstorming, Paper Prototyping, etc.) used by two Belgian companies. Both companies observed are using Paper Prototyping [7], which enables design UIs. Figure 3 on page 4 represents a Paper Prototyping session with a cycle of four different activities.

User Centered Design (UCD) represents a collection of user-centric methods and more generally a philosophy for approaching technology design. UCD methods are not isolated activities of development process. Theses activities are concentrated on the construction of usable and ergonomic UIs. UCD engages the user, its activities and their environment in all stages of an interactive application's analysis and design.

The CUCD method is presented independently of any technological support. However, offering a tool support for the method considering the collaboration among distributed teams requires also to take into account a multi-surface context where designers and final users they could working together using multiple devices to interact.

In the **Drawing**, designers define the structure of the interface, producing some sketches depicting a UI in the form of screens, roughly one for each *"state"* of the interface. This activity is then iteratively performed between the elicitation of requirements and design activities when designers and users need to obtain the system specifications.

Designers and end-users would sometimes be divided into subgroups. An aspect very important to achieve as it facilitates the collaborative sessions between stakeholders and end-users. The second phase covers **Prototyping** where designers defined the **behavior** by assembling the different screens in a logical manner either in a flip-book or in a sequence of screens linked with arrows in order to make the navigation explicit.

The next step is the activities of **Sharing**/**Testing**. At this stage, each subgroup sticks their produced prototype on a wall, arranged in a logical way, forming a sort of storyboard. An overall validation happens at that stage, where the navigation flows are tested. Very often the participants point to the drawings following the previously designed navigation, in a sort of *"Storytelling"*. The last stage is the **Discussing**/**Reflecting** where once the design is discussed and the possible problems are highlighted, the modifications are agreed. Then the group proceeds to another iteration of drawing/prototyping/sharing until both parts are satisfied with the results.

Based on these observations, our collaborative method is divided into four basic phases: *Structure Definition*, *Behavior Definition*, *Test Definition*, and *Reflect Definition*. The goal of devising such method is to support the activities identified on the paper prototyping practice, keeping in mind the basics of Sketching and Prototyping as tools for helping designers and users to communicate, and also to include technological support. From the Drawing activity, we have

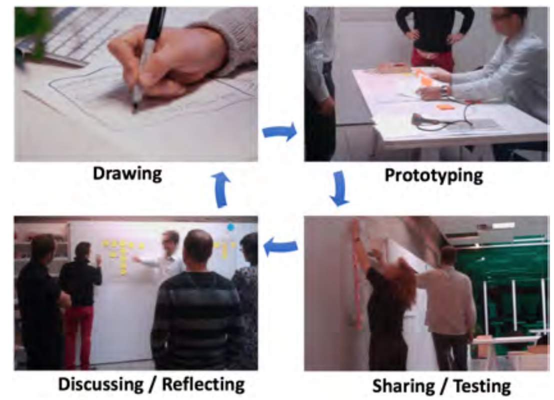

Figure 3: UI design sessions based on Paper Prototyping.

abstracted the Structure Definition, from the Prototyping we have abstracted Behavior Definition, and respectively from Sharing/testing to Test and Discussing/Reflecting to Reflect. The work-flow that links all the disciplines together is presented in Figure 4.

The Structure Definition Phase
Designers define the structure of the interface, depicting a UI from screens either by sketching them or by composing interface elements. Sketches are created according to the gathering requirements. All the elements of the UI are set according to the interaction need of the user and to the functional specifications. In this step, which can be conducted while interviewing the user, the Designer concentrates on the content of the interface and on what is going to be shown.

The activities begin with the definition of the structure by designers and users and iterate **([structure iteration]** / **[structure ready])** proceeding to the next step when both parts agree. Figure 4 details this step further, having the

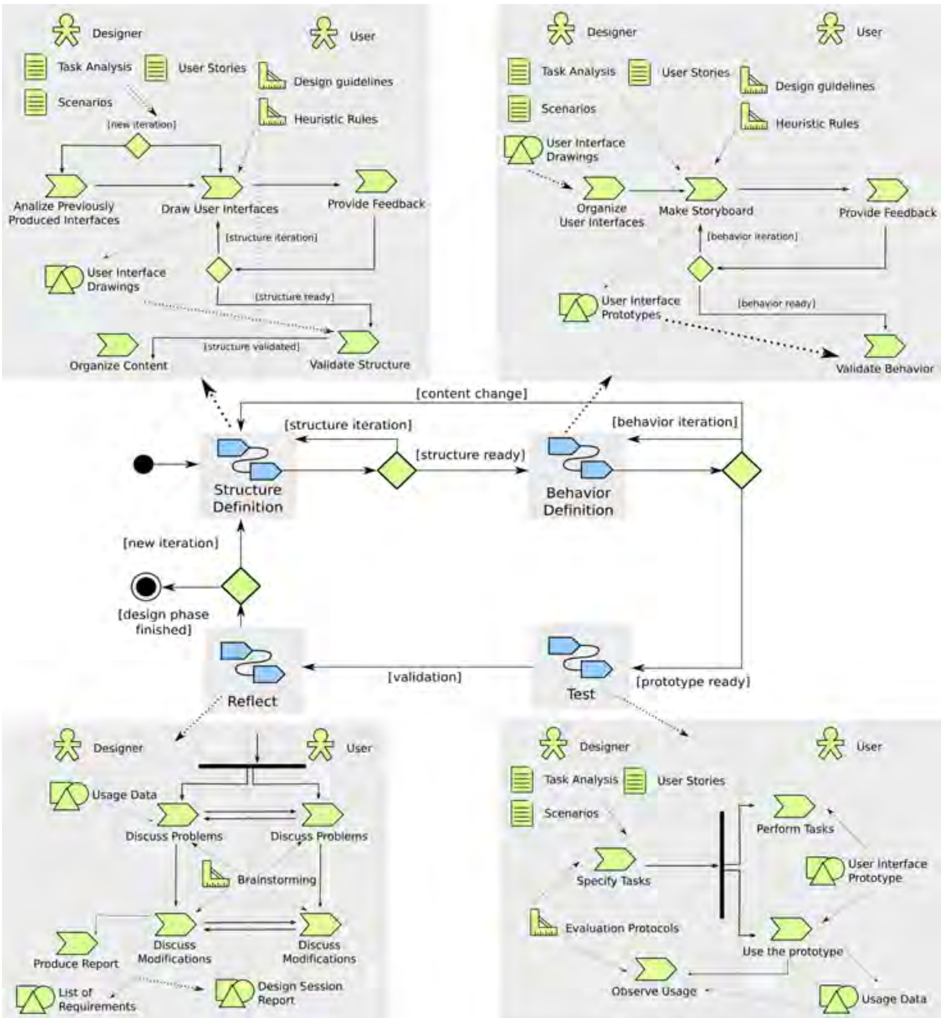

Figure 4: The Method Workflow.

Designer starting the process by either Analyzing Previously Produced Interfaces and Drawing UI or simply by drawing them. For those activities, the Designer uses models such as User Stories, Task Analysis, Scenarios and Guidelines such as Design Guidelines and Heuristic Rules. The User should provide feedback about the produced interfaces until the structure is agreed and validated. Then, the Designer should organize the content in order to proceed to the next step.

The Behavior Definition Phase

In this phase, Designers focuse on the practical usage of the system, or how the information will be shown, while the different screens are assembled in a sequential way (see Figure 4). The activities start with the Designer assembling the different UIs produced previously and then proceed to Make the Story-board which connects the interfaces in a logical manner, taking into consideration the models produced in the Structure Definition. The UI Prototypes are created in this activity for further validation by the User. They might also iterate back to Structure definition prior to proceeding to the next step, since the definition of behavior very often leads to insights about the content that will be displayed.

The Test Definition Phase

In this phase, each subgroup puts their prototypes on the wall, arranged in a logical way, forming a sort of storyboard (see Figure 4). This step requires the active participation of the User, since this is the role that validates the prototype. The Designer specifies the task to be conducted, that can match the Scenarios, User Stories, Task Analysis, etc. to be evaluated, either by planning a user study with Evaluation Protocols. The User should perform the tasks while using the prototype, and the Designer has the role of observing the usage, from which usage data might be collected and evaluated.

The Reflect Definition Phase
Designers and Users can use brainstorming techniques to discuss possible problems and modifications. This activity starts by splitting the work between Designers and Users in parallel. Nevertheless both perform essentially the same activity, since they discuss problems and modifications. Designers have the responsibility to produce a report containing a Requirement List (regardless of whether the overall iteration is over or not). The list can be seen as a series of post-it notes stuck to the wall, (see Figure 3 on page 4). In case the iteration is over, a design session report is produced, otherwise a new iteration can begin.

Discussion
One of the highest priorities in flexible and agile software development process is the development of runnable software that can be validated by both Stakeholders and Users. Generally, a prototype is used as a proof of concept. It can be considered as a first model towards a software development process. The holistic view is an iterative process of discovery, coding and validation, with small iterations.

This vision does not consider that system requirements will all be provided correctly. The proposed method is a recommendation for rapid gathering and validating requirements, for which UCD can be used. The inclusion of the method within UCD and Agile is aligned with previous proposals, such as *Communication-Centered* design and *Extreme Designing* [4], where the focus is on the communication between designers and users. In order to show how to include multiple devices or surfaces on the method, a combination of the configurations can be defined which considers the different users (or roles, which contains both designers and users) and devices. We can consider single and multiple users and also, single and multiple devices, thus giving a combination of 4 possible configurations of roles and devices (see Figure 1). Each step of the method can use one of the 4 possible *"user X device"* combinations, being expressed by n^r, where n is the number of combinations and r is the steps of the method. In total, there are 256 possible combinations of single/multiple users, single/multiple devices and the phases Structure definition, Behavior definition, Test, Reflect.

Figure 5: GAMBIT session on a phone, a tablet and a tabletop computer.

GAMBIT is the acronym for Gatherings And Meetings with Beamers and Interactive Tablets. The main capabilities and features of the tool are:

- Multi-surface and collaboration.
- Infinite workspace.
- Supports the four step of the collaborative UCD.

GAMBIT
GAMBIT is "a multi-surface collaborative tool for UI design that allows the sketching and simulation of UIs on many different devices" [12]. The tool was developed using a web technology - Google Web Toolkit (GWT). Therefore, it is accessible online and works on almost every modern web browser which makes it reachable from various platforms and interaction surfaces. Basically if the device has a sufficiently updated browser application, GAMBIT will run on it. Figure 5 illustrates this point showing a use of gambit on the same session with a phone, a tablet and a tabletop computer.

GAMBIT allows the user to draw on virtual sheets of paper on a wall, that is represented as an infinite workspace (Figure 6). Many devices can be used at the same time as they show parts of that workspace in an analogous way to a map software. Users can zoom in/out to show more or less details, sketch interfaces and "play" them as if they were real systems. In GAMBIT the "Interface production" (see Figure 7) is conducted mainly by designers. It eventually includes users for providing feedback, validating the structure and optionally producing themselves the interface drawings. By activating, the Skectch mode, it is possible for the user to create "scenes" and sketch on them. The designer can add, remove and organize scenes spatially. GAMBIT includes an import feature allows lets the possibility for the user to con-

Figure 6: The infinite workspace concept of the GAMBIT

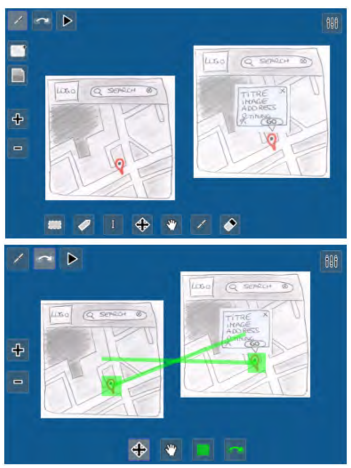

Figure 7: The Interface production (above) and the Behavior Definition (down) of GAMBIT.

struct interfaces in different levels of fidelity: low (sketching), medium (wireframe mockup) and high (UI screenshot).

The "Prototype production" (see Figure 7) is also conducted mainly by designers and eventually includes users for providing feedback, validating the behavior and optionally producing prototypes themselves. With the Prototype mode, the user is able to create interactive zones on top of the scenes (e.g. a "next" button) and connect these zones to another scene existing in the defined structure (e.g. the scene that is targeted by the "next" button).

The last feature consists in the test of the developed prototype in any device. The designer can select the screen where he wants to start and the tool will remove all other interface controls showing only the selected scene in full screen. The prototype can be navigated since the defined interactive zones will react to users' input, showing the next scene when an action occurs.

Experimentation

Table 1 shows acceptable results of a user trial applied to 20 participants having different backgrounds. Participants were divided into 2 groups and instructed to sketch a system UI for children. Participants are asked to fill the IBM Computer Usability Questionnaire (CSUQ) to give their feedback about GAMBIT. The IBM CSUQ was increased with questions concerning the usage of animation pictures, and UI navigation concerns. Group 1 responded that the information provided in GAMBIT is easy to understand (Q13), while group 2 responded that GAMBIT is simple to use (Q2). Regarding the least voted questions Group 1 disagrees that the moving images are not recommended (Q20). Finally, the Group 2 stated that they did not believe that they were productive quickly using GAMBIT (Q8).

	Group 1	Group 2
(1)	13 $(\mu = 6,15; \sigma = 0,67)$	2 $(\mu = 7,00; \sigma = 0,00)$
(2)	20 $(\mu = 4,30; \sigma = 2,05)$	8 $(\mu = 3,25; \sigma = 0,91)$
(3)	6 $(\mu = 4,64; \sigma = 1,04)$	20 $(\mu = 4,40; \sigma = 1,85)$
(4)	2 $(\mu = 6,96; \sigma = 0,20)$	2 $(\mu = 6,04; \sigma = 1,02)$

Table 1: Cumulated results of the CSUQ. (1)=The most suffrage question; (2)=The less suffrage question; (3)=The most critical participant; (4)=The less critical participant.

Conclusion and future works

We have presented a CUCD method based on sketching to support early prototyping along flexible and agile software development process. The approach is far from introducing new software development methodology that integrates UI engineering with the process, but is rather an attempt to include the steps of the proposed method into existing practices through User Centered Design. The method is presented as independent of technological support, which is then added in order to support the assessed requirements. GAMBIT, constructed to support the method, takes into consideration the contemporary technological support of multiple-surface. Its tools have the original property of functioning with multiple possible combinations of designers, users and devices at the same time. Future works will focus on describing how the UIs designed by sketches with our CUCD method can be linked with others activities that happen within a regular software development process. In this view, we consider performing several experimental studies which involve classical known SE methodologies to find new insights and improvements to our method.

Acknowledgement

The authors would like to thank Ugo Sangiorgi for the big contribution to the GAMBIT tool.

REFERENCES

1. AMBLER, S. W. Agile adoption rate survey results: March 2007, 2007.
2. BELLAMY, R., DESMOND, M., MARTINO, J., MATCHEN, P., OSSHER, H., RICHARDS, J., AND SWART, C. Sketching tools for ideation (nier track). In *Proc. of the 33rd Int. Conf. on Software Engineering* (New York, NY, USA, 2011), ICSE '11, ACM, pp. 808–811.
3. BUXTON, B. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007.
4. DA SILVA, B. S., AURELIANO, V. C. O., AND BARBOSA, S. D. J. Extreme designing: Binding sketching to an interaction model in a streamlined hci design approach. In *Proceedings of VII Brazilian Symposium on Human Factors in Computing Systems* (New York, NY, USA, 2006), IHC '06, ACM, pp. 101–109.
5. FOX, D., SILLITO, J., AND MAURER, F. Agile methods and user-centered design: How these two methodologies are being successfully integrated in industry. In *Agile, 2008. AGILE '08. Conference* (Aug 2008), pp. 63–72.
6. GONZALEZ-PEREZ, C. Filling the voids - from requirements to deployment with open/metis. In *ICSOFT 2010 - Proceedings of the Fifth Int. Conf. on Software and Data Technologies, Volume 1, Athens, Greece, July 22-24, 2010* (2010), p. 19.
7. LANCASTER, A. Paper prototyping: The fast and easy way to design and refine user interfaces. *Professional Communication, IEEE Transactions on 47*, 4 (Dec 2004), 335–336.
8. LANDAY, J. A., AND MYERS, B. A. Sketching interfaces: Toward more human interface design. *Computer 34*, 3 (2001), 56–64.
9. MCINERNEY, P., AND MAURER, F. Ucd in agile projects: Dream team or odd couple? *interactions 12*, 6 (Nov. 2005), 19–23.
10. MESZAROS, G., AND ASTON, J. Adding usability testing to an agile project. In *AGILE* (2006), J. Chao, M. Cohn, F. Maurer, H. Sharp, and J. Shore, Eds., IEEE Computer Society, pp. 289–294.
11. OSSHER, H., VAN DER HOEK, A., STOREY, M.-A., GRUNDY, J., AND BELLAMY, R. Flexible modeling tools (flexitools2010). In *Proceedings of the 32Nd ACM/IEEE Int. Conf. on Software Engineering - Volume 2* (New York, NY, USA, 2010), ICSE '10, ACM, pp. 441–442.
12. SANGIORGI, U. B., BEUVENS, F., AND VANDERDONCKT, J. User interface design by collaborative sketching. In *Proc. of the Designing Interactive Systems Conference* (2012), ACM, pp. 378–387.
13. VAN DER LUGT, R. Functions of sketching in design idea generation meetings. In *Proc. of the 4th Conference on Creativity & Cognition* (New York, NY, USA, 2002), C&C '02, ACM, pp. 72–79.
14. WHITTLE, J., HUTCHINSON, J., ROUNCEFIELD, M., BURDEN, H., AND HELDAL, R. *Industrial adoption of model-driven engineering: are the tools really the problem?* LNCS. Springer, 2013, pp. 1–17.
15. WUEST, D., SEYFF, N., AND GLINZ, M. Sketching and notation creation with flexisketch team: Evaluating a new means for collaborative requirements elicitation. In *Requirements Engineering Conference (RE), 2015 IEEE 23rd Int.* (Aug 2015), pp. 186–195.
16. YANG, M. C. Observations on concept generation and sketching in engineering design. *Research in Engineering Design 20*, 1 (2009), 1–11.

Built-In Device Orientation Sensors for Ad-Hoc Pairing and Spatial Awareness

Jens Emil Grønbæk
Aarhus University
Aarhus, Denmark
jensemil@cs.au.dk

Kenton O'Hara
Microsoft Research
Cambridge, United Kingdom
keohar@microsoft.com

Abstract
Mobile devices are equipped with multiple sensors. The ubiquity of these sensors is key in their ability to support in-the-wild application and use. Building on the ubiquity we look at how we can use this existing sensing infrastructure combined with user mediation to support ad-hoc sharing with nearby devices. In particular, the paper proposes a technique for exploiting the built-in compass in mobile devices to aid the process of pairing them for ad-hoc sharing in a variety of proxemic arrangements and F-formations.

Author Keywords
Mobile Devices; Spatial Awareness; Proxemics; F-formations

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous; See [http://acm.org/about/class/1998/]: for full list of ACM classifiers. This section is required.

Introduction
Research prototypes on cross-device interaction enable real-time collocated sharing between devices on the web [3] and through spatial awareness of other devices [7, 4]. However, there remains a large challenge in designing such systems with a view to enabling ad-hoc social interactions - that is, systems that do not require assumptions about prior online relationships or systems that do not rely

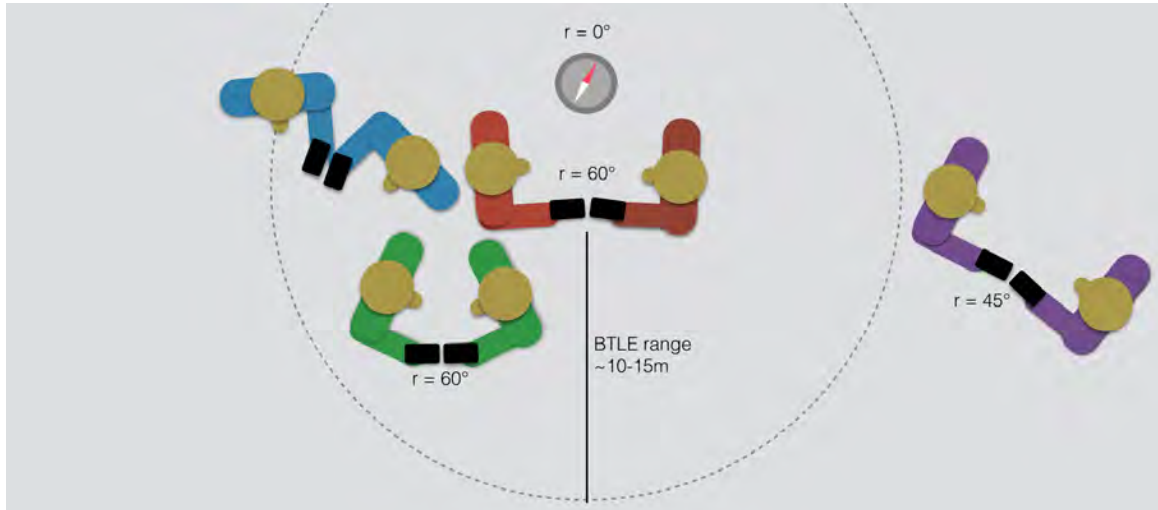

Figure 1: Sensor data to simplify device pairing and ad-hoc sharing in the wild. A sensor fusion technique using orientation data together with other inputs can provide a filtering algorithm for detecting devices.

on additional external instrumentation with sensors. State-of-the-art mobile technology has the potential of making pairing and sharing between collocated devices an easy and lightweight experience by utilizing sensor capabilities. Commercial apps make use of e.g. NFC [1], local networks [2], or by bumping devices together [3]. However, most available approaches still have limited cross-platform support.

In response to these challenges, we explore the potential of built-in orientation sensors in the design of a cross-platform interaction technique for simple setup and reconfiguration of the spatial relationship between devices. The paper contributes with three ideas for cross-device interactions: 1) an interaction technique for ad-hoc pairing of devices using web technologies and orientation sensors, 2) a preliminary outline of trade-offs to address with limited coarse-grained sensor information, and 3) a position on how knowledge of proxemics and F-formations can inform design of mobile interactions.

Point2Share focuses on ad-hoc sharing between devices. It uses orientation and distance between people as a way of spatially organising your sharing relationships. The interaction techniques are inspired by typical patterns in F-formations, such as face-to-face, corner-to-corner and side-by-side formations [2, 4] (see figure 1). It is inspired by Marquardt et al's explorations on F-formations and spatial awareness in GroupTogether [4] and takes off from their future work motivation that more easily deployable sensing modalities and inherent trade-offs should be investigated. The proposed techniques address two challenges for using

[1] https://developer.android.com/training/beam-files/index.html
[2] https://support.apple.com/en-us/HT203106
[3] http://bu.mp/

built-in mobile sensors to enable in-the-wild use: 1) insufficient knowledge of devices' spatial relations, and 2) balancing user and sensor input in design of interaction techniques. Before presenting Point2Share in more detail, we first provide a review of related work that helps situate our current system and its particular rationale, in the broader landscape of other systems.

Related Work

It has been suggested that knowledge of people's spatial relationships, or *proxemics*, based on sensor input, can be used for mediating interactions. *Proxemic interactions* is envisioned as devices with fine-grained knowledge of nearby people and other devices to enable more natural social interactions [1, 4]. However, in state-of-the-art mobile technology the available sensor information is much more coarse-grained than what can be provided in a lab with dedicated sensor technology. A range of pairing techniques and interaction techniques using sensors and spatial awareness have been proposed in HCI as well as in products. In the following, these are compared with their ability to scale for in-the-wild use.

Pairing Techniques for Collocated Sharing

In the current state of the mobile market, apps that make collocated sharing easier have moved from the apps in the mobile app stores into the operating system with examples such as Bump (acquired by Google), Android Beam and AirDrop. Common for all three is their exploitation of existing built-in mobile sensors. Bump enables one-to-one sharing by bumping phones together and detects collision with the accelerometer. Android Beam uses NFC and is a very compelling technology for ad-hoc sharing operating at Android system level. However, this is currently limited to Android devices. AirDrop also allows for wireless sharing between devices through Wifi and Bluetooth, however, it is limited to the Apple ecosystem of devices. Furthermore, Webstrates has been proposed as an experimental system with a flexible model and compelling applications for real-time collocated sharing. Webstrates are web-based substrates that can contain content, computation and interaction possibilities [3]. Pairing in Webstrates can be as simple as sharing a URL. However, sharing relations require prior setup, and this can be simplified and become more dynamic by the use of sensors.

Spatial Awareness

The notion of F-formations is a theoretical construct that articulates patterns of spatial configuration of people in social interaction. [2]. F-formations have been used by Marquardt et al to infer certain patterns from sensors to have system awareness of how a group of people orient towards each other and devices. This concept was explored through the GroupTogether system [4]. Although GroupTogether uses relatively lightweight sensor instrumentation of an environment, such a requirement nevertheless limits the more ubiquitous application of these ideas beyond that instrumented space. Other HCI systems have explored the fine-grained spatial awareness between devices. HuddleLamp [7] enables mobile devices to have spatial awareness of each other, enabling simple spatially intuitive content sharing between collocated devices. However, as with GroupTogether, there is a dependence upon additional external instrumentation that restricts its more ubiquitous use: HuddleLamp requires cameras in the surrounding environment and mobility of devices is limited to near the table surface.

In the wild – with no assumptions of instrumentation in the environment – the granularity of spatial awareness in the above-mentioned systems is not available. However, pinching has been suggested as a way of stitching displays together [6].This technique is very intriguing for in-the-wild

Technical Problems

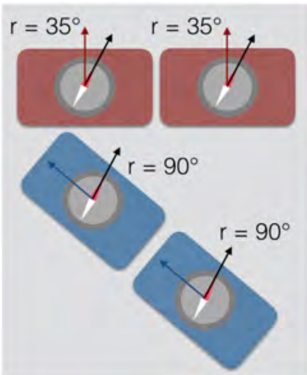

Figure 2: Correlating absolute orientation as implicit parameter for easy pairing.

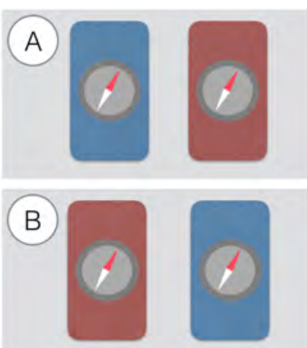

Figure 3: The problem of inferring relative orientation. There is no way from the sensor data to infer the difference between blue device being to the left (A) or to the right (B).

applications. What should be noted about this interaction is that it requires devices to be in contact in order to make the connection. For some social situations, this might not always be a proper way of interacting. And furthermore, this technique requires fully manual recalibration when users move a device.

Point2Share: Interaction Techniques

We developed Point2Share to demonstrate how the above-mentioned issues could be overcome in future systems. The system is web-based to enable use on any mobile platform. Point2Share enables users to share content by transferring from one device to another in a way that the two displays extend each other. The system uses synchronised web clients powered by the Webstrates infratructure [3]. Content could be anything from work-related documents or a youtube video to contact information, or even permissions and control units for remote controlling another client. Point2Share currently involves two interaction techniques that are based on orientation awareness: *pairing* and *re-configuration of spatial sharing relation*.

Filtering Algorithm for Ad-Hoc Pairing

In order to have the two displays of the mobile devices extend each other, they must first be paired. Pairing is based on the assumption that orientation is an indicator of intention [1, 4]. Users can think of their mobile device as a pointer pointing towards the target they want to pair with. To generate a match we propose a sensor fusion technique combining multiple implicit parameters for pairing. A single pairing data point contains *orientation* information from the compass, *proximity* information from Bluetooth and pairing requests from within a limited *time window*. A server can use these data points to pair two devices and distinguish their pairing from others. Assuming that two users are holding their phone in ordinary portrait mode in their hands, the absolute orientation of two devices can be used to infer that they are pointing towards each other. Use of orientation data is illustrated in figure 2 and in (video) figure 4. However, in the wild, situations might occur where multiple people are pairing with similar sensor data points. As an example illustrated in figure 1, the red pair can be distinguished from the purple in both absolute orientation and proximity, but red and green have similar orientation data pairs and are all in proximity of each other. If they are also trying to pair within the same time window, there is no way to fully infer the correct match. However, in this case a filtered list of possibilities will be presented to all devices, and the users have to choose between a list of people. When the two users have selected each other from the filtered lists, they are assigned a shared webstrate for real-time sharing. Bottom line is that when sensor data is insufficient, the user will be prompted to make the decision from presented information.

Reconfiguring Relative Orientation

In the pairing we assume that the top edges of the device are pointing towards each other. We cannot be sure from sensor data because of the problem illustrated in figure 3, but it can be a fair assumption when in the pairing process. However, to support greater flexibility in the spatial relation between devices after pairing, we can allow for reconfiguring the spatial device relations with a quick recalibration. Since this *reconfiguration* mode is separated from the *pairing* mode, the following technique can assume that this is only about relating two devices to each other.

In Ballendat et al's conceptual framework for Proxemic Interactions they mention *orientation* as a dimension of proxemics [1]. They distinguish between absolute and relative orientation. What we want is to have some notion of relative orientation between two people, but we only have absolute

Initial Prototypes

https://youtu.be/1vZZzPd6KnE

Figure 4: Pairing using orientation as part of the input to a filtering algorithm. This short video figure demonstrates how absolute orientation can be used as input for pairing.

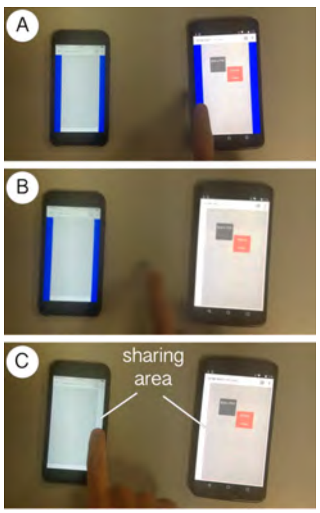

Figure 5: Setup of side-by-side configuration. Two devices are detected as being side-by-side due to similar orientation values (A). A tap on the Android device on its left edge (B) and a tap on the iPhone on its right edge (C) sets up a natural sharing relation where Alice and Bob can share content with spatial cues

orientation (i.e., a phone's orientation relative to North). The final step of going from absolute to relative orientation between two devices is to have knowledge of which side the other device is on relative to your device. With a single tap on the edge (blue edges in figure 5), the user provides the input indicating which direction the sharing relation points towards. Limiting to two facing formations between devices – side-by-side and face-to-face depicted in figure 7 – allows for a spatial flexibility in that the switch between the two modes only happen at a certain threshold. Thus, it does not restrict the devices to be in a very particular relation to each other, but rather just within the semi-circle that is within the limits of the 180-degree threshold.

Transferring Content Between Devices

Pairing sets up a spatial relationship between two devices. Users can then share between devices by *dragging* between them, where one device display appears as an extension of the other, like the *portals* technique in GroupTogether [4]. This is demonstrated in figure 6. It is currently implemented in a prototype, where two users each have a web client viewing different parts of a shared canvas. They have virtual documents that are only visible on each their visible part and can drag between them. By dragging content to the white area at one edge of the display, a user sends a document in this direction towards the other paired device. Sharing a canvas is susceptible to privacy issues, however, this could be implemented in a way that would keep users' content outside of the shared canvas until they intend to share.

Point2Share: Applications

While the design space for Point2Share is yet to be fully explored, we outline initial example applications to illustrate what Point2Share would enable for ad-hoc sharing. The flexibility of web-based infrastructures such as Webstrates

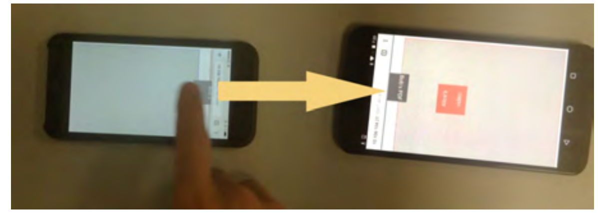

Figure 6: Transferring with spatial awareness between devices, as in portals [4].

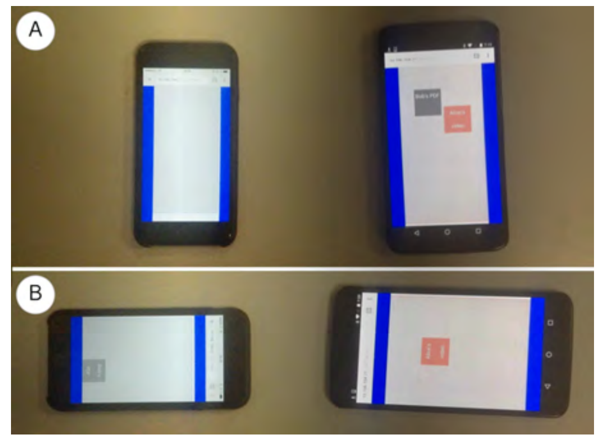

Figure 7: Two supported F-formations: side-by-side (A) and face-to-face (B).

[3] combined with built-in mobile sensors enables applications for a variety of different proxemic arrangements.

One-to-One Relationships

As outlined in the prototype, the pairing can happen between two devices and easy reconfiguration can allow for arranging in a variety of F-formations. In the demo application, two people can share between their devices by dragging a digital document towards the direction of the other person's device in physical space. The interaction of pointing towards another device provides an embodied way of sharing an object. This would allow for a natural way of sharing contact information or business cards in a face-to-face formation. Objects for sharing could also take other forms, like permissions or control elements, similar to the proposals in [3]. In a side-by-side (or corner-to-corner) formation, users might be able to stitch their device displays together. This could enable sharing relations like looking through a merged photo album with photos from the same vacation, or a merged list of suggestions based on two people's Netflix profiles like proposed in [8].

One-to-Many Relationships
The prototype from this paper explores one-to-one relations. However, it is not hard to imagine extending this to one-to-many relationships. A different kind of proxemic arrangement could be in group scenarios, where users can quickly switch between sharing to a public display or their neighbor. This would require a setup where users have calibrated their relative orientation to the public display by once pointing towards it from their position in a meeting room or auditorium. Afterwards they would be able to point in that direction and drag content to the shared screen. This would provide quick reconfiguration of proxemic relations ranging from sharing in a more intimate/personal space with the person next to you or in a more social/public space with a group. A drawback of this approach is that it requires recalibration when a user moves.

Discussion

Current State of Web-based Sensor APIs
An important point about the implementation is that it is done with web technology. For mobile interactions in the wild it is desirable that different mobile platforms are able to communicate. Fortunately, the device orientation data is available from within a web application. The device orientation API is available for web applications. Currently, however, Bluetooth for proximity sensing is not part of the implementation and would still in status quo have to be implemented in a native mobile app. However, the future is pointing towards Bluetooth coming to web apps [4], and there are currently tools for easy development of cross-platform Bluetooth support [5].

Proxemic Interactions and Interaction Proxemics
Situating the Point2Share system alongside other proximity-based sharing apps offers an opportunity for some reflections on proxemics as they relate to interaction. If we consider Beam, Bump and JuxtaPinch we can see how these entail a particular socio-spatial relationship between people that is bound up in the spatial characteristics of the pairing and exchange mechanisms. All three of these systems require the devices to be physically close or touch in order for sharing to occur. These proxemic characteristics of the devices then entail the users of the devices to also be physically close. This is intriguing in the sense that permissions for sharing are dealt with implicitly – that is, such deliberate physical closeness reflects a certain social familiarity or intimacy between the sharers that implies the trust necessary for content exchange.

However, these proxemic requirements of the system also limit some of the spatial flexibility under which collocated exchange might happen. The Point2Share system, while acknowledging the social benefits of systems requiring physical closeness, seeks to offer more spatial flexibility while still maintaining a sensibility to proxemic relations between devices and people as outlined in the concept of Interaction Proxemics [5]. It does so by not making systems

[4] https://developers.google.com/web/updates/2015/07/interact-with-ble-devices-on-the-web?hl=en
[5] https://github.com/randdusing/cordova-plugin-bluetoothle

decisions purely on basis of implicit parameters. Rather the pairing/filtering technique of Point2Share addresses a key issue with using imperfect contextual information for implicit interactions by proposing a suggestive method that balances user and system input for informed decision making. Users are prompted to make decisions together with the sensor data to make the simplest possible pairing process, provided with only limited sensor data. This is different from most of the techniques proposed on proxemic interactions, e.g., [1, 4]. In this way, the system affords greater flexibility in the ways that people can configure themselves with respect to each other to enact their social relationships. At the same time it allows these people to exploit certain proxemic characteristics and configurations to make the ad hoc collocated sharing of content simple and more intuitive.

Conclusion and Future Work

This paper takes an initial step in exploring orientation awareness using the compass for social device interactions. Initial prototypes have confirmed that it is possible to use the compass for comparing device orientations for pairing and distinguishing between two spatial device configurations. The sensor data is coarse-grained and thus puts a limit on the decisions we can make on the basis of it. A full prototype with proximity detection needs to be built to test the robustness of the proposed interaction technique with the heterogeneity of built-in mobile sensors.

REFERENCES

1. Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic Interaction: Designing for a Proximity and Orientation-aware Environment. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. ACM, 121–130.

2. Adam Kendon. 2010. Spacing and Orientation in Co-present Interaction. In *Development of Multimodal Interfaces: Active Listening and Synchrony*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–15.

3. Clemens N Klokmose, James R Eagan, Siemen Baader, Wendy Mackay, and Michel Beaudouin-Lafon. 2015. Webstrates: Shareable Dynamic Media. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 280–290.

4. Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device Interaction via Micro-mobility and F-formations. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, 13–22.

5. Helena M Mentis, Kenton O'Hara, Abigail Sellen, and Rikin Trivedi. 2012. Interaction proxemics and image use in neurosurgery. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. ACM Press, New York, New York, USA, 927.

6. Heidi Selmer Nielsen, Marius Pallisgaard Olsen, Mikael B. Skov, and Jesper Kjeldskov. 2014. JuxtaPinch: Exploring Multi-device Interaction in Collocated Photo Sharing. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '14)*. 183–192.

7. Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. HuddleLamp: Spatially-aware mobile displays for ad-hoc around-the-table collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 45–54.

8. Henrik Sørensen and Jesper Kjeldskov. 2014. *Proxemic interactions with multi-artifact systems*. . . . Journal on Advances in Intelligent Systems.

Towards Cross-Surface Immersion Using Low Cost Multi-Sensory Output Cues to Support Proxemics and Kinesics Across Heterogeneous Systems

Rajiv Khadka
University of Wyoming, 3DIA Lab
1000 E. University Ave,
Dept. 3315, Laramie, WY 82071
rkhadka@uwyo.edu

James Money
Idaho National Laboratory
Applied Visualization Laboratory
995 University Blvd.,
Idaho Falls, ID 83401
James.Money@inl.gov

Amy Banic
University of Wyoming, 3DIA Lab
Idaho National Laboratory, AVL
1000 E. University Ave,
Dept. 3315, Laramie, WY 82071
abanic@cs.uwyo.edu

Presented at the Cross-Surface '16 workshop, in conjunction with ACM ISS'16.
November 6, Niagara Falls, Canada.

Abstract
Collaboration in immersive systems can be achieved by using an immersive display system (i.e. CAVE and Head-Mounted Display), but how do we communicate immersion cross-surface for low immersive displays, such as desktops, tablets, and smartphones? In this paper, we present a discussion of proxemics and kinesics to support based on observation of physical collaboration. We present our research agenda to investigate low-cost multi-sensory output cues to communicate proxemics and kinesics aspects cross-surface. Doing so may increase the level of presence, co-presence, and immersion, and improve the effectiveness of collaboration cross-surface.

Author Keywords
Proxemics, kinesics, virtual environment, collaboration, immersion, multi-sensory output cues

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Figure 1: Example of an Immersive System (CAVE Automated Virtual Environment).

Figure 2: Example of a low immersive system.

Figure 3: Example of a non-immersive system.

Introduction

Collaborative Virtual Environments (CVEs) provides a shared virtual workspace where people from different geographical locations can meet and interact with each other. They can share and work on 3D objects/models to achieve common goals [3]. CVEs are increasingly being used to support collaborative work between co-located or geographically remote collaborators. Collaborators can work between different levels of immersive display systems, from highly immersive (CAVE and Head-Mounted Display) to low immersive (desktop, tablet, and smartphone) display systems [1]. However, there are difficulties in providing a realistic user experience and immersive quality to users in low immersive systems when collaborating with others who are interacting in an immersive environment in a shared networked virtual environment. Immersive display systems have a high degree of presence in comparison with a display system that does not support aspects to create immersion [1]. Increased sense of presence and co-presence in a system helps to increase the effectiveness of the collaborative work [4], transfer of skills to real world [12], and improves skills in collaborative manipulation [15]. Only a few have conducted design and evaluation of collaboration using cross-display collaborative virtual environments [19]. CVEs for collaboration provide a strong platform for learning, understanding and evaluating complex spatial information. While working in CVEs, a decrease in immersion due to the field of view, the field of regard, display size, tracking available, etc. leads to a decrease in co-presence. Since co-presence has been shown to enhance collaboration performance [5], we would like to explore alternative methods to increase co-presence among systems with lower levels of immersion. In this paper, we present aspects in proxemics and kinesics which should be supported cross-surface due to importance for effective collaboration. We present our research agenda for designing and investigating low-cost multi-sensory output cues to be used for a low immersive display system in order to communicate proxemics and kinesics from immersive systems to non-immersive systems. Our goal is to design and investigate low-cost proxemics cue techniques as a way to increase co-presence such that these cues help to improve collaborative workflow across heterogeneous systems for a variety of applications.

Proxemics

Proxemics refers to the study of how space and distance influence communication [11]. There are four larger zones of proxemics: public, social, personal, and intimate distance. The amount of space defined for each of these zones varies across cultures and social organizations. Public and social zones refer to space that is four or more feet away from our body, and the communication that typically occurs in these zones is formal and not intimate [11]. Communication that occurs in the social zone, which is four to twelve feet away from our body, is typically in the context of a professional or casual interaction, but not intimate or public. Personal and intimate zones refer to the space that starts at our physical body and extends a short distance (for US Americans it is about four feet). These zones are reserved for friends, close acquaintances, and significant others. The intimate zone, reserved for only the closest friends, family, and romantic/intimate partners, extends within about one to two feet of the body. With individuals entering this space, it is difficult for individuals to ignore others in this space. One additional smaller zone of proxemics is pericutaneous space [7]. This refers to a layer of space very near to

Figure 4: Collaboration in an Immersive Display System

Figure 5: Collaboration in a non-immersive display system. Example shown is video avatar projected in the space.

and a feeling just prior to touching. Visual-tactile perceptive fields overlap in processing this space. For example, an individual might see a feather as not touching their skin but still experience the sensation of being tickled when it hovers just above their hand. Other examples include the blowing of wind, gusts of air, and the passage of heat. We are very interested in studying this space in the context of cross-surface interaction. Using this phenomenon, we plan to add feedback mechanisms to serve as cues to indicate nearness, distance, direction of human contact.

Kinesics

There are three main types of gestures: adaptors, emblems, and illustrators [11]. Adaptors are touching behaviors and movements that indicate internal states typically related to arousal or anxiety. Emblems are gestures that have a specific agreed-on meaning. Illustrators are the most common type of gesture and are used to illustrate the verbal message they accompany.

Advances in Collaborative Virtual Environments

A highly immersive room-sized, fully dynamic real-time 3d scene capture and continuous-viewpoint head-tracked display on a life-sized tiled display wall was developed for symmetric collaboration (i.e. using same immersive display system across all remote participants) [17]. The system used abundant sensor information, cameras, high-quality displays for increased level of immersion. Another research work supporting symmetric collaboration incorporated three-way communication over a distributed shared workspace which was designed to support three channels of communication: person, reference, and task-space [20]. Researchers developed a highly immersive telepresence system for symmetric collaboration that allowed distributed group of collaborators to meet and share their 3D environment among each other easily [2]. More recent work involving collaboration across heterogeneous devices, such as supporting interaction across 3D workstation, desktops, and Personal Digital Assistants (PDA's), was conducted [20]. A data-centric design was used for synchronous collaboration. A recent approach for collaboration used large interactive virtual spaces i.e. high-resolution wall-sized displays and CAVE for remote collaboration across heterogeneous collaborative virtual environment [9]. This research identified that sense of presence in a virtual environment helps in understanding users' interaction capabilities which eventually helps to increase the workflow of the collaborative task. Other important aspects for collaboration are co-presence, social presence, spatial proximity, relationships, and context, which have all been shown to range from higher to lower for face-to-face interaction, immersive CVEs, video-CVEs, and standard CVEs respectively [14,18]. The face-to-face condition was experienced as being significantly helpful in collaborative task [13]. Proxemics have effects on the collaborator's behavior, communication, and social interactions. Proxemics also enhance the collaborator's ability to work together and collaborate easily. [8,10]. So, with higher proximity and more cues of awareness, we can increase copresence and which then should increase efficiency in collaboration tasks.

Proxemics and Kinetics Found in Physical Collaboration with Single-Display

We conducted a study to investigate how users collaborate for data exploration and analysis in the

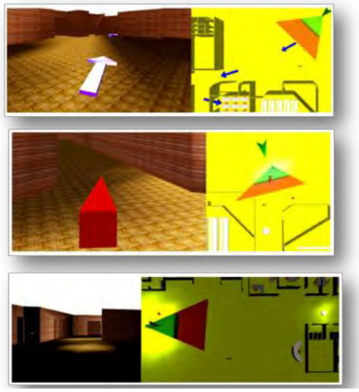

Figure 6: Visual Cues for proxemics. Photo from [6]

Figure 7: Participant wearing vibrotactile belt in their arms using low immersive display system during collaboration.

same physical environment or co-located space (see fig. 1), using an immersive display system at Idaho National Laboratory. We collected action and observation data on participants exploring data environments and performing analysis tasks while collaborating with other participants in the co-located space. This study was designed to help us understand how the participants would interact and use the immersive virtual environment while collaborating their work. Our analysis of the data collected describes the collaborative behavior and interactions exhibited by the participants in a co-located immersive space while interacting with an immersive application.

Proxemics to Support

While working with remote collaborative virtual environments, sense of presence between collaborators can be preserved better through 3-dimensional video avatars [4]. However, current implementations do not necessarily preserve all the relationships between the user and environment as they would appear to be in the physical space of the immersive display. These types of relationships should be preserved when providing these representations. In the next section we outline our plan to help increase the user experience of the collaborators in relation to spatial direction, distance, and awareness of type of collaborator. Furthermore, there should be a distinction between touching, nearness to touch, within reach but not touching, and not in reach. The reason is that effective collaboration is correlated with the combination of the factors nearness and relationship of collaborators. Use of audio as interaction techniques may also help to increase the level of user experience collaborators experiences. Spatial audio should be preserved cross-surface to facilitate collaborative communication. Spatial audio helps with the challenge of multiple collaborators talking in a CVE.

Kinesics to Support

Gestures that supported collaboration were broken down into most frequent and specific gestures used. A majority of the participants were trying to grab and touch data being displayed in the application. The next most frequently used gestures were those which included pointing at a specific area of the immersive application in which they want to collaborate and communicate with the co-located collaborators. This result shows that spatial relationships between users and each other, users and data, and users and the display are essential for communication and task completion for collaboration. So, when we are using heterogeneous collaborative virtual environment which includes low immersive display systems communicating the proximity play an essential part for effective collaboration. Spatial relationships of gestures between users, between users and environment, and between users and the display, need to be preserved and rescaled to map appropriately to the physical space of each individual collaborator. Representations need to be provided for the users' gestures, however, can be enhanced through visual interaction history trails to better illustrate the physical body to environment relationship. We preserve these relationships by providing cues in the environment.

Research Agenda for Designing and Investigating Proxemics and Kinesics

In this section, we present our position on designing and investigating low-cost proxemic cues techniques for low immersive display systems. We have identified that there is a need to add feedback mechanism to serve as cues to indicate nearness, distance, and direction from human contact. There is also a need to communicate gesture action, direction, and content. In this section, we discuss and present our research agenda to design and investigate proxemic and kinesic cues to enhance collaboration cross-surface.

Levels of Proxemics

In our prior work we found that known collaborators, the distance is closer for more effective collaboration. For collaborators who do not know each other well, further distance makes for more effective collaboration. We seek to answer the question, what if we could break these conventions in cross-surface? For example, for known collaborators always communicate nearness, even when far away. For unknown collaborators, communicate in public space, even if in the virtual environment their locations are rather close. There are multiple levels of proxemics, as defined in the section on 'Proxemics'. We shall investigate the use of visual, vibrotactile feedback, and spatial audio to determine which and what variations more appropriately communicate nearness, touch potential, distance, direction, and collaborator relationship from an immersive environment to a non-immersive environment. In the following sections, we detail more specifically what we will investigate for each.

Visual Cues

Based on data from our user study, using natural gestures is the most effective method of interaction in an immersive virtual environment. Spatial relationships of gestures between users, between users and data, and between users and the display, need to be preserved and rescaled to map appropriately to the physical space of each individual collaborator. So, we want to use visual cues i.e. glyphs (to be used as a control condition) and a light source (see fig. 5) with projectors to provide and increase the sense of presence in the appropriate location while using a low immersive display system. A set of visual solutions we will investigate include cues on the edges of the display frame, small projector-based solutions in the environment, and glassware with small LEDs in the periphery. Aspects to communicate include varying the type of representation for collaborator relationship, the size and/or color based on proximity, and location based on direction. To communicate pericutaneous space, we will need to investigate subtle ubiquitous changes in visual information. These visual solutions may help individual collaborators to communicate their directional and distance information in respect to another collaborator more effectively during collaboration.

Glyphs: Each collaborator will be assigned an individual glyph. Each glyph will have a distinct color and a name of collaborator on top of the glyph. So, that each glyph can be distinguished easily. With the increase in the nearness of the collaborator, the size of the glyph will also be increased and otherwise. With the increase in the nearness of the collaborator, the distance of the glyph will also decrease. The glyph will be pointed to show the point of view of the collaborator.

Small Pico Projectors: In low resolution, each collaborator will be assigned a separate color of light source. For high resolution, the avatar of the collaborator will appear. With the increase in the nearness of collaborator, the intensity of the light will also increase and decrease as the distance of the collaborator increases. The light source will be projected depending upon the direction of the position of the other collaborator.

Glassware with LEDs: The intensity of blinking will increase or decrease with the increase or decrease in the distance of the collaborators. The left led or right led will blink in respect to a drastic change in the direction of the collaborator. A constant glow will emit based on the direction of the collaborator and slightly increase or decrease in intensity based on distance.

Vibrotactile Cues

There has been an increasing amount of work using vibrotactile belts in the virtual environment to communicate a sense of presence and touch. We also plan to investigate the use vibrotactile belts, shoulder and arm-wear, or within the device itself. Collaborators to wear vibrotactile belts on both arms (see fig. 6). One solution is that when a collaborator is moving near to the collaborator, the vibration of the belt will increase or decrease in frequency or strength accordingly to inform the collaborator the direction and distance in respect to the another collaborator. Low vibration or locational vibration may distinguish between pericutaneous space and less intimate space. These vibrotactile cues may help to increase the sense of presence and copresence while a collaborator is using a low immersive display system and will be compared to our visual solutions.

Temperature and Pressure Cues

We will also explore the production and release of temperature and pressure as cues. We will vary the increase and decrease as well as the frequency applied.

Spatial Audio Cues

Spatial audio should be preserved across the immersive visualization environments to facilitate collaborative communication. However, in this section we discuss in addition to spatial audio itself, to use as a means to communicate other proxemics and kinesics information. It will be a challenge to be able to balance any use of audio with actual audio from collaborators. These may be in the form of non-verbal sounds or more ubiquitous rising and falling of consistent music in the background. We will compare the differences among visual, vibrotactile and spatial audio cues.

Summary and Conclusion

In this paper, our research agenda to design and investigate low-cost multi-sensory output cues as a way to increase co-presence such that these cues help to improve collaborative workflow cross-surface from immersive systems to non-immersive systems. Spatial relationships between users and environment, users, and users with the display technology are important and need to be preserved across heterogeneous environments among collaborators whether those are high immersive or low immersive display system. And the preservation of interaction, gestures, communication mechanisms, and spatial relationships should be adapted cross-surface for simulating immersion.

References

1. Axelsson, Ann-Sofie, Åsa Abelin, Ilona Heldal, Ralph Schroeder, and Josef Wideström. "Cubes in the cube: A comparison of a puzzle-solving task in a virtual and a real environment." CyberPsychology & Behavior 4, no. 2 (2001): 279-286.
2. Beck, S., Kunert, A., Kulik, A., & Froehlich, B. (2013). Immersive group-to-group telepresence.

Visualization and Computer Graphics, IEEE Transactions on, 19(4), 616-625.

3. G. Bente, S. Rüggenberg, N. C. Krämer, and F. Eschenburg, "Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations," Human Communication Research, vol. 34, no. 2, 2008.
4. Bulu, Saniye Tugba. "Place presence, social presence, co-presence, and satisfaction in virtual worlds." Computers & Education 58, no. 1 ,2012.
5. Bystrom, Karl-Erik, Woodrow Barfield, and Claudia Hendrix. "A conceptual model of the sense of presence in virtual environments." Presence: Teleoperators and Virtual Environments 8, no. 2 (1999): 241-244.
6. Duval, Thierry, Thi Thuong Huyen Nguyen, Cédric Fleury, Alain Chauffaut, Georges Dumont, and Valérie Gouranton. "Improving awareness for 3D virtual collaboration by embedding the features of users' physical environments and by augmenting interaction tools with cognitive feedback cues." Journal on Multimodal User Interfaces 8, no. 2 (2014): 187-197.
7. Elias, L.J., M.S., Saucier, (2006) Neuropsychology: Clinical and Experimental Foundations. Boston, MA. Pearson Ed. Inc.
8. Engleberg, Isa N. (2006). Working in Groups: Communication Principles and Strategies. My Communication Kit Series, pp140-1
9. Fleury, Cédric, Nicolas Ferey, Jean-Marc Vézien, and Patrick Bourdot. "Remote collaboration across heterogeneous large interactive spaces." In Collaborative Virtual Environments (3DCVE), 2015 IEEE Second VR International Workshop on, pp. 9-10. IEEE, 2015.
10. Hall, E.T. (1966) The Hidden Dimension. Anchor.
11. Hans, Anjali and Hans, Emmanuel. Kinesics, "Haptics and Proxemics: Aspects of Non -Verbal Communication", Iosr journal of humanities and social science (iosr-jhss) volume 20, **issue** 2, ver. iv (feb. 2015), pp 47-52 e-issn: 2279-0837.
12. Harris, Helen, Jeremy N. Bailenson, Alexia Nielsen, and Nick Yee. "The evolution of social behavior over time in second life." Presence: Teleoperators and Virtual Environments 18, no. 6 (2009): 434-8.
13. Harrison, B., H. Ishii & M.H. Chignell (1994). An empirical study of orientation of shared workspaces and interpersonal spaces in video-mediated collaboration. Telepresence Technical Report OTP-94-2. Ontario Telepresence Project.
14. Hauber, Jörg. (2008). Understanding Remote Collaboration in Video Collaborative Virtual Environments. Diss. University of Canterbury.
15. Held, Richard M., and Nathaniel I. Durlach. "Telepresence." Presence: Teleoperators & Virtual Environments 1, no. 1 (1992): 109-112.
16. J. Jung, H Park, D. Hwang, M. Son, D. Beck, J. Park; Woojin, "A Review on Interaction Techniques in Virtual Environments" in Proceedings of the International Conference on Industrial Engine;2014, p1582
17. Maimone, Andrew, and Henry Fuchs. "A first look at a telepresence system with room-sized real-time 3d capture and life-sized tracked display wall." Proceedings of ICAT 2011, to appear (2011): 4-9.
18. Short, J., Williams,E., and Christie, B.. "The social psychology of telecommunications." (1976).
19. Skarbez, Richard, and Mary C. Whitton. "Enabling Distributed Collaboration among Heterogeneous Devices." (2009).
20. Tang, Anthony, et al. "Three's company: understanding communication channels in three-way distributed collaboration." Proceedings of the 2010 ACM conference on Computer supported cooperative work. ACM, 2010.

Designing Multi-Surface Environments to Support Collaborative Sensemaking

Leila Homaeian
Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada
leila.homaeian@uwaterloo.ca

James R. Wallace
School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada
james.wallace@uwaterloo.ca

Stacey D. Scott
Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada
School of Computer Science, University of Guelph, Guelph, ON, Canada
stacey.scott@uwaterloo.ca

Presented at the Cross-Surface '16 workshop, in conjunction with ACM ISS'16.
November 6, Niagara Falls, Canada.

Abstract
In a prior project, we investigated multi-surface data browsing techniques to support collaborative sense-making, specifically involving geospatial data. During this work, we studied two interactive techniques that facilitated data browsing across a digital tabletop and personal tablets. The study findings indicated that one cross-surface data browsing technique afforded the use of *Least Collaborative Effort (LCE)* among team members during their collaborative interactions. LCE is a phenomenon that occurs during human discourse whereby people proactively assist each other while communicating to minimize the overall effort exerted by a group, allowing efficient communication. This paper presents a new doctoral project focused on employing the social principle of LCE to enhance the design—and interaction efficiency—in multi-user, multi-surface environments.

Author Keywords
CSCW, least collaborative effort; multi-surface interaction; sensemaking; data exploration.

ACM Classification Keywords
H.5.m. Information interfaces and presentation.

Introduction

Analysis and decision-making involving large and/or complex data sets requires significant cognitive effort, and typically involves multiple cognitive stages. Experts need to overview, comprehend, and interpret the data to arrive at informed decisions, using in a process known as sensemaking [14,15]. Sensemaking is often conducted collaboratively to help bring in various perspectives to help identify key facts, filter out irrelevant information, establish connections within the data, and develop deep insights [5,13]. Large surfaces (e.g. digital tabletops) and multi-surface environments involving large surfaces (e.g. tabletop plus tablet environments) have been shown to facilitate the collaborative sensemaking process by providing a large shared reference point in which to review and discuss relevant data [14].

In an ongoing project, we have developed a multi-surface environment to explore different interaction techniques that facilitate collaborative sensemaking. In a recently completed case-study [3], we explored cross-surface data browsing techniques designed to support sensemaking involving geotagged data; that is, data that has a direct correlation to a geographic location. In a comparative laboratory study of two different cross-surface data browsing techniques, we discovered, quite unexpectedly, that one technique—a technique that allowed flexible "ownership" of the data connection between the shared digital tabletop and team members' personal tablets—allowed team members to assist each other in cross-surface data browsing activities during periods of the overall task involving joint analysis and discussion. This behaviour helped minimize the expected difficulties that participants might have experienced with the data browsing technique, including potential reach issues that may have occurred on the digital tabletop due to the direct-touch nature of the technique. Instead, we observed group members appropriate their collaborators' tools in a manner that benefitted the overall joint goal of the group.

This observed cooperative behaviour was consistent with behaviour observed in human communication studies, known as the *Least Collaborative Effort* (LCE) phenomenon [2], whereby discussants help each other during communication exchanges to minimize the overall effort exerted during joint discussions. In LCE behaviour, for instance, when someone completes another's sentence, it helps people communicate more efficiently and effectively.

The goal of this research is to leverage people's social tendency to exhibit LCE behaviour during collaborative interactions to improve the efficiency and effectiveness of collaborative computer interactions within multi-surface environments. This work will be conducted as part of the first author's dissertation work.

Applying Social Theories to Surface Technology Design

This work aims to apply the knowledge and understanding of LCE social behaviour from the field of communications theory to the design of collaborative multi-surface environments. This design approach builds on previous work, within the surface computing and the broader HCI field, of applying social theories to the design of surface technologies. This section overviews some of the prior work, namely in the area of applying social theories relating to territoriality, proxemics, and communication grounding.

Researchers have applied knowledge of human territoriality practices to the design of digital tabletops [7,12], wall displays [6], multi-display environments [11], and virtual environments [8]. This work applies the understanding that implicit and explicit organization of shared environments into areas with different levels of accessibility, primarily personal and group spaces, can provide certain individual cognitive benefits (e.g. by providing space for individuals to disengage from the group activity to concentrate on an independent task activity) and certain collaborative benefits (e.g. by minimizing communications needed to share task resources). Multi-surface environments that provide a large shared display typically enable territorial behaviour by providing both shared and personal "territories" for group members to use during different phases of their joint tasks.

Researchers have also applied knowledge of proxemics behaviour to the design of large surfaces [1,4] and multi-surface environments [9]. This work applies the understanding that humans exhibit specific social behaviours, and have different levels of psychological comfort, at different physical distances. Researchers have applied this theory to design technologies that attempt to respond dynamically and "appropriately" to people and their devices when they are at different distances that naturally afford different social behaviours, or in the case of devices, interaction behaviours.

Our research aims to draw similar human-computer interaction inspiration from knowledge of the LCE developed in the human communications field to intentionally design multi-surface sensemaking environments that allow people to perform cooperative actions within the digital environment that help to minimize the overall group effort. For example, in many multi-user environments each collaborator must perform similar, parallel actions to view the same data on their own personal displays. Designing with a view to support LCE, the environment may instead allow someone to proactively help their collaborator(s) view a common data. Before discussing this notion further, we first provide some background on our recent work in the area of collaborative multi-surface sensemaking environments, and evidence from a recent study that revealed people's tendency to exhibit LCE behaviour when provided with the right tools.

Our Prior Work

Prior research on collaborative sensemaking indicates that groups commonly fluctuate between periods of independent and jointly-coupled work in a process known as *mixed-focus collaboration* [13]. This behaviour allows group members to efficiently investigate the available data, e.g. by dividing up the data browsing task, independently reflect on the data and potential hypotheses under consideration about the data, and collaboratively discuss and debate the data and potential insights it provides. Yet, there is a lack of a research focused on designing environments that allows group members to switch between individual and collective aspects of their work as needed.

In prior work [14], we showed that digital tabletops, and a multi-surface environment that provides a shared digital tabletop and personal tablets were both effective in supporting the collaborative sensemaking process. An advantage of the multi-surface environment is that it provides separate—but connected—spaces for the different independent and joint work that occurs during mixed-focused collaboration.

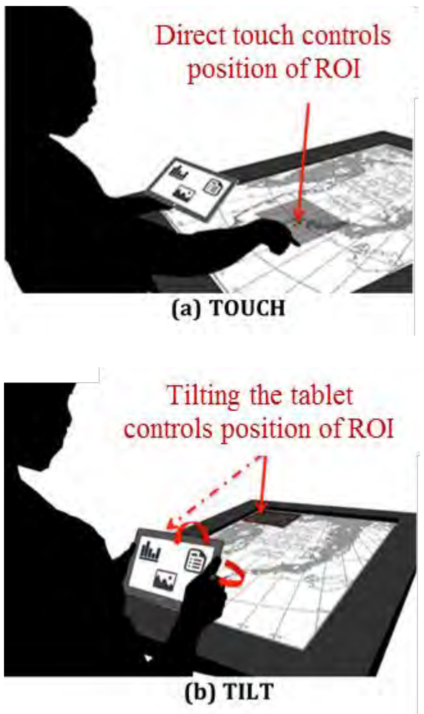

Figure 1. Multi-surface data browsing interaction: a) Using a direct TOUCH gesture on the tabletop to control the ROI movement, and b) Using a TILT gesture on the tablet to control the ROI movement on the tabletop.

Following this work, we began to explicitly investigate the design of multi-surface environments that support the different phases and overall process of collaborative sensemaking, including the type of mixed-focused collaboration inherent to this task. Our initial research focus involved the design of cross-surface data browsing techniques to support efficient independent overview (i.e. getting an overall sense of the available data), which is commonly reported as a key first step in the collaborative sensemaking process [15].

As a case study, our research group developed and evaluated several cross-surface data browsing interaction techniques within the context of a geospatial sensemaking task [3]. That is, the data set of interest involved digital information that was associated with specific geographic locations. An experimental multi-surface prototype environment was first developed that included a large (65-inch) multitouch tabletop and personal multitouch tablets. The tabletop showed a map of a certain geographic region with data icons overlaid on the map at specific locations with any associated data (e.g. charts, graphs, images, informational bulletins, and images). The tablets allowed each group member to view data associated with the data icons within a certain 'Region of Interest' (ROI) selected on the tabletop map. The design drew from existing Overview-plus-Detail (O+D) visualization work, which provides two separate interfaces for overview and detailed views [10], where the tabletop provided the "Overview" and the tablets provided the "Detail" view of available data.

Multi-Surface Interaction Mechanisms

Two interaction techniques, namely TOUCH and TILT (Figure 1), allow users to select data icons on the map and then view the associated detail on their tablets. TOUCH employs the multi-touch technology of the digital tabletop. It allows users to directly touch the ROI on the map and drag it to a region of interest. In contrast, TILT utilizes the built-in device sensors in the tablet. Thus, it involves a 3-D tilt motion above the tabletop to move the ROI to a new location. In TILT, pressing a 'hold' button on the tablet while tilting the device moves the ROI on the tabletop. TOUCH has flexible ownership, meaning team members can help each other move the ROI using this mechanism. This can encourage cooperation among parties. In TILT, on the other hand, the movement of the ROI can only be controlled by the person wishing to view the associated data; they don't need to interrupt their partner when the region of interest cannot be physically reached.

Key Study Findings

Full findings from the comparative study of the TOUCH and TILT techniques can be found in Goyal [3]. Here we summarize key findings related to this research. It was expected that reachability and independent exploration offered by the TILT technique would be advantageous in a tabletop-centric environment. Interestingly, however, the study revealed that the flexible "ownership" of the TOUCH technique let the participants use a wider variety of collaborative strategies to cooperate as a team. Relying on their partners for moving the ROI encouraged more collaboration; often one team member was responsible for moving the ROI for the group.

The Theory of Least Collaborative Effort

The above observed behaviour is consistent with the theory of LCE. This is in spite the fact that we did not intentionally design our tool to support LCE behaviour. When exhibiting this behaviour, "participants try to minimize their collaborative effort—the work that both do from the initiation of each [communication] contribution to its mutual acceptance" [2, p. 226]. The theory of LCE relates to a broader communication theory known as *communication grounding* [2]. Grounding is a collection of mutual beliefs, knowledge, and assumptions essential for communication between two or more people. It ensures that the people communicating are "on the same page" and clearly understand each other's mutual beliefs so that they can carry on an effective conversation. Grounding is established through an iterative process of clarification.

This iterative process is where the LCE phenomenon comes into play. While attempting to carry on a conversation, people frequently perform various assistive behaviours to help minimize the overall effort that all parties involved have to exert to continue moving forward with the conversation. For instance, a listener may finish a speaker's sentence to indicate that they understand what the speaker is saying. This prevents the speaker from having to fully articulate their meaning. The same listener might instead utter a "yeah, yeah, yeah" along with a head nod to indicate that they understand what the speaker is saying, making it unnecessary for the speaker to provide further explanation. Both of these behaviours—a relevant next turn, or verbal and non-verbal acknowledgments—provide *evidence* to the speaker that the listener acknowledges (i.e. understands) the intended contribution. Overall, people engaged in a communication process try to minimize their collective effort –or exhibit LCE—to achieve common ground.

Current Research Questions

The shared visual space in our current environment helps team members know and/or anticipate what knowledge everybody else has. This knowledge, of course, is limited to what is displayed on the tabletop. We intend to employ the LCE theory to enhance the design of our sensemaking environment so that people have a deeper understanding of the current cognitive stage that their team members are at. More specifically, we are studying how a multi-surface environment can enable LCE behaviour and thereby better support grounding.

According to the grounding concept, if a collaborator knows what knowledge their partner has, the communication becomes more effective. Therefore, we will explore different data sharing mechanism to allow team members see what their partner is viewing or has previously viewed. The intent is to improve mutual understanding about what the other party knows, and consequently support establishment of mutual belief.

Moreover, evidence of mutual understanding of the data being explored is essential in moving the communication forward. To this end, a feature letting partners acknowledge that they have received what has been shared is necessary.

Consider the following scenario: Two team members, called *A* and B, are each exploring data associated with different region of interests (ROIs) on their personal tablets. *A* finds some insightful information and decides to share it with *B*. They could drag the data and drop it on the tabletop so that both can view and discuss it. This approach resembles working around physical tables and therefore it provides an intuitive data sharing mechanism. However, the data may obstruct some parts of the map and interfere with the overall sensemaking. Alternatively, *A* could press a button and send the data to *B*'s tablet. Now both can view the same data without having to reposition themselves around the tabletop or occupying its surface. Yet, this may interfere with *B*'s individual work.

We plan to explore alternative multi-surface design options for data sharing with the goal of improving a group's ability to exhibit LCE behaviour, and to ultimately foster grounding without compromising other aspects of sensemaking.

Conclusion

This paper discusses designing multi-surface environments that facilitate collaborative sensemaking around complex data. We discuss how well-established theories from the social sciences have been applied by researchers to design surface technologies. In particular, we are interested in the communication grounding and Least Collaborative Effort theories, which have been used to understand how people communicate effectively and efficiently. In a new project, we plan to explore multi-surface design solutions that employ these theories to facilitate effective collaborative sensemaking. At the workshop, we hope to share our work with the community and connect with other surface computing researchers interested in applying theories from social sciences to designing more effective digital tools.

Acknowledgments

This work is partially funded by the Natural Sciences and Research Council of Canada (NSERC). We also thank Nippun Goyal for his contributions to the earlier work discussed in this paper.

References

[1] Cheung, V. and Scott, S.D., 2015. Studying Attraction Power in Proxemics-Based Visual Concepts for Large Public Interactive Displays. *Proc. ITS 2015*, 93-102.

[2] Clark, H.H. and Brennan, S.E., 1993. Grounding in Communication. *Readings in Groupware and Computer-Supported Cooperative Work*, R.M. Baecker Ed. Morgan Kaufmann, 222-234.

[3] Goyal, N., 2016. *Investigating Data Exploration Techniques Involving Map Based Geotagged Data in a Collaborative Sensemaking Environment.* Master's Thesis, University of Waterloo, Waterloo, ON, Canada.

[4] Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M., 2011. Proxemic interactions: the new ubicomp? *interactions* 18(1), 42-50.

[5] Isenberg, P., Fisher, D., Paul, S.A., Morris, M.R., Inkpen, K., and Czerwinski, M., 2012. Co-Located Collaborative Visual Analytics around a Tabletop Display. *IEEE Trans. on Visualization and Computer Graphics,* 18(5), 689-702.

[6] Kister, U., Reipschl, P., Matulic, F., and Dachselt, R., 2015. BodyLenses: Embodied Magic Lenses and Personal Territories for Wall Displays. *Proc. ITS 2015*, 117-126.

[7] Klinkhammer, D., Nitsche, M., Specht, M., and Reiterer, H., 2011. Adaptive personal territories for co-located tabletop interaction in a museum setting. *Proc. ITS* 2011, 107-110.

[8] Kunert, A., Kulik, A., Beck, S., and Froehlich, B., 2014. Photoportals: shared references in space and time. *Proc. CSCW* 2014*; social computing*, 1388-1399.

[9] Marquardt, N., Hinckley, K., and Greenberg, S., 2012. Cross-device interaction via micro-mobility and f-formations. *Proc UIST* 2012, 13-22.

[10] Plaisant, C., Carr, D., and Shneiderman, B., 1995. Image-browser taxonomy and guidelines for designers. *IEEE Software 12*, 2, 21-32.

[11] Ramos, G., Hinckley, K., Wilson, A., and Sarin, R., 2009. Synchronous Gestures in Multi-Display Environments. *Human–Computer Interaction 24*, 117-169.

[12] Scott, S.D. and Carpendale, S., 2010. Theory of Tabletop Territoriality. In *Tabletops - Horizontal Interactive Displays*, C. MÜLLER-TOMFELDE Ed. Springer, 375-406.

[13] Tang, A., Tory, M., Po, B., Neumann, P., and Carpendale, S., 2006. Collaborative coupling over tabletop displays. Proc *CHI*2006, 1181-1190

[14] Wallace, J.R., Scott, S.D., and MacGregor, C.G., 2013. Collaborative sensemaking on a digital tabletop and personal tablets: prioritization, comparisons, and tableaux. *Proc CHI* 2013, 3345-3354

[15] Yi, J.S., Kang, Y.-a., Stasko, J.T., and Jacko, J.A., 2008. Understanding and characterizing insights: how do people gain insights using information visualization? *Proc BELIV 2008*, 1-6.

Considering Collaboration in *ʔeləw̓k̓ʷ — Belongings*

Reese Muntean
250-13450 102nd Avenue
Surrey, BC V3T 0A3 Canada
rmuntean@sfu.ca

Presented at the Cross-Surface ’16 workshop, in conjunction with ACM ISS’16.
November 6, Niagara Falls, Canada.

Abstract

Bringing multiple interactive surfaces and smaller devices together is possible, but not yet extensively explored in the HCI community generally nor in the context of museum spaces specifically. While this is an exciting area for research, collaboration utilizing interactive surfaces as well as collaboration across multiple devices requires careful consideration when incorporated into the design of some museum installations. In this position paper, I introduce the context for the development and design of *ʔeləw̓k̓ʷ — Belongings*, an interactive tangible tabletop installed in a museum and designed to communicate Indigenous traditional knowledge and cultural values. I discuss some of the specific design decisions made around collaboration among museum visitors, the use of multiple devices, and the calculated decisions to exclude certain aspects collaboration and interaction. This discussion highlights our considerations of both previous design research as well as complex colonial histories.

Author Keywords

Tangible interaction; intangible cultural heritage; digital heritage; Museum of Anthropology; Musqueam Indian Band; c̓əsnaʔəm.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces, Evaluation/Methodology.

Introduction

Bringing multiple interactive surfaces and smaller devices together is possible, but not yet extensively explored in the HCI community generally nor in the context of museum spaces specifically. While this is an exciting area for research, collaboration utilizing interactive surfaces as well as collaboration across multiple devices requires careful consideration when incorporated into the design of some museum installations.

ʔeləw̓kʷ — Belongings is one example of an interactive tangible table in a museum setting. The table was developed by a team of researchers, designers, and curators from Simon Fraser University and the University of British Columbia (UBC) in Vancouver, Canada. Installed in the Museum of Anthropology (MOA) at UBC for the *c̓əsnaʔəm, the city before the city* exhibition, this tabletop was designed to communicate the traditional knowledge and cultural values of the Musqueam Indian Band. Visitors use replicas of ancient belongings excavated from c̓əsnaʔəm and everyday objects in contemporary Musqueam lives to interact with the table and learn stories of the Musqueam community's past and how their culture and knowledge continues today.

In this position paper, I will illustrate why cooperation around interactive surfaces requires careful consideration in certain museum settings through my experience as project manager of *ʔeləw̓kʷ — Belongings*. First, I will introduce the context for the development and design of *ʔeləw̓kʷ — Belongings*. This context is important in understanding the design decisions made around the way museum visitors would experience collaboration and the inclusion of multiple devices in the installation. I will then discuss some of the specific design decisions made around collaboration among museum visitors, the use of multiple devices, and the calculated decisions to exclude certain approaches to collaboration and interaction. This discussion will highlight the development team's considerations of both previous design research as well as complex colonial histories.

Context

c̓əsnaʔəm is an ancient village site of the Musqueam people, located near the mouth of the Fraser River in what is now Vancouver, British Columbia. Archaeological evidence suggests people lived in this location for over three thousand years, and according to Musqueam oral history, their ancestors have lived there from time immemorial [23].

c̓əsnaʔəm, the city before the city is a partnership between the Musqueam Indian Band, the Museum of Vancouver, and MOA, along with the University of Waterloo. In this series of three exhibitions, the institutions introduce visitors to c̓əsnaʔəm, each with a different focus.

ʔeləw̓kʷ — Belongings was part of the MOA exhibition (which ran from January 2015 to January 2016) in which MOA curators Susan Rowley and Jordan Wilson were attempting to challenge the meaning of an archaeology exhibit. Visitors would likely expect to see ancient belongings (commonly referred to as "objects" or "artifacts") supplemented by interpretations from academic experts and scientific views. The curators wanted to show, though, that material culture is not equivalent to culture; there is much more to Indigenous communities than artifacts. So rather than showing ancient belongings, the MOA exhibit focused on quotes and stories from Musqueam community members, noting that academics and scientists are not the only voices with valuable information to offer.

ʔeləw̓k̓ʷ – Belongings was developed by researchers and curators from UBC and SFU, along with Musqueam input and approval by way of the exhibit advisory committee for *c̓əsnaʔəm, the city before the city*. We designed the interactive tangible tabletop with these curatorial priorities in mind, using twelve replicas of ancient and modern belongings to access these stories around which the exhibition was built, reconnecting tangible and intangible heritage.

To briefly explain the main interaction sequence with the table, each belonging has four categories of information. Visitors place a belonging into an activator ring to start exploring, and each category is accessed through a different set of interactions with the table. When a visitor successfully completes the interactions for each category, information will appear on the table in the form of text, contemporary images, historical documents, and quotes from community members. By completing these interaction sequences, visitors ultimately unlock videos that appear on monitors situated next to the table. These are stories from Musqueam community members about the process of learning, the passing on of traditional knowledge, and their culture.

Given the context of *ʔeləw̓k̓ʷ — Belongings*, we had specific considerations to address when making design decisions. Early in the design process, we articulated nine main goals, including highlighting Musqueam voices, encouraging sharing and conversation among visitors, decolonizing museum spaces. Next I will discuss how these goals influenced decisions made around collaboration among museum visitors, the use of multiple devices, and the calculated decisions to exclude certain aspects of collaboration and interaction.

Collaboration Among Visitors

We were influenced by the work of Antle (who was part of the *ʔeləw̓k̓ʷ — Belongings* development team) et al. conducted research on a tabletop application called *Futura*. A multi-player simulation game to raise awareness of the complexities around sustainable development, *Futura* required players to work together, each taking on a role in charge of a specific resource, and the players would receive feedback as to the positive or negative effect their choices were making on the overall environment. *Futura* was successful in encoding both learning and collaboration within the game mechanics and interface design [1]. Antle continued work on another sustainability game called *Youtopia*, conducting a number of studies examining how to design tangible systems to support collaboration and reflection around personal values (e.g. [2, 3, 7]).

Drawing on Antle’s previous work, we designed an “activation” tool that was mentioned in the previous section, a ring that was necessary to access digital information on the table. Visitors could handle the belongings and set them down on the table, but they would have to use one of the two activator rings to move beyond that initial stage on information.

We recognize that having only two rings would limit the number of visitors directly interacting with the table at any given time. This was partially necessary, though, due to the size of the tabletop. With twelve belongings available for visitors, the tool would limit which ones would bring up information of the surface. Limiting one of the necessary tools would encourage a sharing of resources among visitors and could spark discussions. This sharing of resources was also related to the sharing of information either by watching or discussing with others who were using the table, such as how to use the tabletop or what interactions were necessary.

Incorporating Multiple Devices

Giaccardi and Palen examined how different media and technologies could be combined and how these intersections could support interactions among communities, spaces, and artifacts. Multiple interactive technologies can work together to open up new ways for users to experience and think about heritage and cultural knowledge. The ways that users interact with these technologies can influence how they understand the socially produced meanings and values ascribed to artifacts [9]. Indeed, there are a number of examples of collaborations with Indigenous communities on multimodal projects that allow for archiving, storytelling, and interacting with cultural heritage (e.g. [14, 24]).

In designing *ʔeləw̓k̓ʷ — Belongings,* we carefully considered what media to include, as well as what devices. Ultimately we included text and images, with photographs, sketches, historic documents, text quotes, and videos. While we did discuss having our interactive surface recognize and interact with any phone, we eventually decided to use the interactive tangible table with three wall monitors nearby.

The basic interactions with the tabletop were mentioned in the context section. Each time a visitor accessed information about a belonging, they would see a mix of photos (many taking by community members on their smart phones), historical images or documents, and quotes from community members. Lastly, if visitors completed the series of interactions, a video clip of a community member would appear on a monitor next to the table. These varied types of information were included to convey that Musqueam culture is not of the past, that the community, their traditional knowledge, and cultural values still remain today.

External monitors were included to further our goals of highlighting Musqueam voices while encouraging collaboration among visitors working on the table. One monitor showed a series of fish cutting images, helping to drawing people into the space and give them basic visual clues as to one of the important Musqueam practices. The other two monitors were paired to each of the activator ring tools. We chose to play videos on the external monitors for a number of reasons. Visitors who were near the exhibit, but not physically interacting, could watch and listen. When multiple groups of visitors were working on the table and one accessed a video, it would not prevent or impede another visitor from their exploration (rather than a video clip appearing on the tabletop surface and "pausing" the table, in effect).

We also discussed the option of allowing visitors to interact with the table using their own smart phones, potentially setting their phones on the table as part of the interactions with the table or creating opportunities for visitors to comment or contribute. Due to tight deadlines and technical concerns, we decided against having the tabletop recognize visitors' own phones.

As for having visitors leave textual comments, tweets, etc., we did want to encourage discussion among visitors. One of our goals was to encourage discussion around the cultural values visitors were learning about through their interactions with the system and around their reflections on their own values in relationship to Musqueam culture. Yet the exhibition overall focused on highlighting Musqueam voices, so we ultimately decided against this feedback mechanism. I will elaborate on importance of Musqueam voices in the next section.

Reconsidering Collaboration

In some cases, we purposely excluded forms of collaboration, such as the above example of allowing visitors to instantly comment on or update the tabletop exhibit. Certainly the museums encouraged the use of the exhibit hashtag in social media during *c̓əsnaʔəm, the city before the city*, but there was no sort of instant dialog in the museum space in this interactive tabletop. This was in keeping with our focus on Musqueam voices and is rooted in efforts to decolonize museum practices.

When presenting *ʔeləw̓k̓ʷ — Belongings,* people often make connections to theories of participatory design and co-creation of experience. While we certainly see the relationship, it is important to note the differences and distinctions. Participatory design was developed to address issues with workplace information technology by having designers work directly with those who used the technology in question [4]. Iversen and others later described the need to revitalize the focus on values [15]. Participatory design is now sometimes part of the development process of museum exhibitions, involving audiences in both the design process [26] and designing for visitors' co-creation of experiences within the exhibitions [16, 25].

ʔeləw̓k̓ʷ — Belongings, rather, is an example of working towards the decolonization of museum exhibitions. The late twentieth century North American museology began working towards building new relationships with First Peoples. In 1992 the Assembly of First Nations and the Canadian Museums Association worked together on the development of the Task Force Report on Museums and First Peoples [12], in an attempt to repair the relationships between Canadian institutions and First Peoples and to move towards more open partnerships.

At the same time, digital imaging, databases, and search technologies were rapidly advancing, allowing for museums and First Nations to implement such open partnerships [21] by together developing new tools for both reassembling dispersed collections and creating new forms of access to Indigenous cultural heritage. Canadian institutions and First Nations were now working as design partners [19, 20].

Along with better access to Indigenous cultural heritage, these new partnerships also worked towards bringing fuller representations of heritage into museum exhibitions [17]. Moving beyond a focus on objects, museums and communities are working together to reconnect the intangible forms of knowledge and traditions with these physical belongings in the museum collections and to share this knowledge with museum visitors (e.g. [10, 11, 13, 18, 24]).

While this idea of collaboration may be viewed as largely positive, some scholars warn that inherent asymmetries of power likely still exist in these contexts. In Clifford's 1997 work *Routes: Travel and Translation in the Late Twentieth Century*, he writes about museums as contact zones, drawing on the idea that a contact zone is "the space of colonial encounters, the space in which peoples geographically and historically separated come into contact with each other and establish ongoing relations" [22]. The conditions in these contact zones often include that of inequality, conflict, and coercion, and Clifford saw this consultation process as that of reworking and revisiting the relationship between the museum and the Indigenous people [6].

Some scholars have also warned against the neocolonial potential in the "museum as contact zone" concept. Boast, though he writes from a supportive stance of collaborations between institutions and Indigenous communities, explores the inherent asymmetry of power in these relationships that is often overlooked [5]. Museums must go beyond consultation and cultural sensitivity and incorporate the actual sharing of authority. Drawing on Geismar, Boast notes that the power still remains with those who have the property and capital and who also have the power to display [8].

These complex issues are very much a part of the history between the Musqueam Indian Band and MOA, and were an important factor for us to consider throughout the design process. Wilson, who is also a member of the Musqueam Indian Band, writes on the power relations in the context of the planning and development of *c̓əsnaʔəm, the city before the city*:

"To my mind, this approach is a result of the agency and authority of the advisory group members; as curators or 'researchers' we simply respected their leaderships and depth of knowledge. In other words, we did not afford them power as much as they commanded it by refusing to be reduced or undervalued as informants or suppliers of data. Simultaneously, this gathered together approach is also a result of the community steering the overall exhibit development' this process was established to adhere to the Musqueam community protocols, of which listening to Elders is critical. The approach we used–gathering together and listening to Elders—is as much a result of Musqueam asserting its agency as it is the exhibit partners (MOV, MOA, and the University of Waterloo) relinquishing power and decision-making capacities" [27].

Not only does this discussion serve to highlight the importance of Musqueam voices in the exhibit, but also to show why collaboration by museum visitors may not always be appropriate. The MOA exhibit and *ʔeləw̓k̓ʷ — Belongings*, strive to move beyond the model of consultation and cultural sensitivity by shifting the voice of authority back to the Musqueam people.

Conclusion

In this paper I described the context of *ʔeləw̓k̓ʷ — Belongings,* an interactive tangible tabletop designed within a specific context, for an exhibition in a heritage institution and in a relationship working towards repairing a long colonial history. I discuss design decisions made in regards to collaboration utilizing interactive surfaces as well as collaboration across multiple devices. And while this context may be quite specific, it serves as an example of instances in which seemingly benign desires to foster collaboration among users of an interactive surface may need further consideration and reflection.

Museums continue to develop exhibits and installations incorporating interactive surfaces, multi-touch tables, and tangible interfaces. There are certainly great opportunities both in museum settings and beyond to explore how multiple technologies or devices might be combined and how collaboration could be better supported, though in some instances there may be types of collaboration and methods used to encourage such collaboration that need to be carefully considered.

Acknowledgements
Special thanks to the Musqueam Indian Band, Museum of Anthropology at the University of British Columbia, Maia Jensen at the SIAT Solid Space Fabrication Lab, and Grand NCE.

References
[1] Antle, A. N., Bevans, A., Tanenbaum, J., Seaborn, K., & Wang, S. Futura: design for collaborative learning and game play on a multi-touch digital tabletop. In *Proc. TEI 2011*, ACM Press (2011), 93–100.

[2] Antle, A. N., Warren, J. L., May, A., Fan, M., & Wise, A. F. Emergent dialogue: eliciting values during children's collaboration with a tabletop game for change. In *Proc. IDC 2014*, ACM Press (2014), 37–46.

[3] Antle, A. N., Wise, A. F., Hall, A., Nowroozi, S., Tan, P., Warren, J., … Fan, M. Youtopia: a collaborative, tangible, multi-touch, sustainability learning activity. In *Proc. IDC 2013*, ACM Press (2013), 565–568.

[4] Björgvinsson, E., Ehn, P., & Hillgren, P.-A. Participatory Design and "Democratizing Innovation." In *Proc. PDC 2010*, ACM Press (2010), 41–50.

[5] Boast, R. Neocolonial Collaboration: Museum as Contact Zone Revisited. *Museum Anthropology*, *34*(1), (2011), 56–70.

[6] Clifford, J. *Routes: Travel and translation in the late twentieth century*. Harvard University Press, 1997.

[7] Fan, M., Antle, A. N., Neustaedter, C., & Wise, A. F. Exploring How a Co-dependent Tangible Tool Design Supports Collaboration in a Tabletop Activity. In *Proc. GROUP 2014*, ACM Press (2014), 81–90.

[8] Geismar, H. Cultural property, museums, and the Pacific: reframing the debates. *International Journal of Cultural Property*, *15*(2), (2008), 109–122.

[9] Giaccardi, E., & Palen, L. The Social Production of Heritage through Cross-media Interaction: Making Place for Place-making. *International Journal of Heritage Studies*, *14*(3), (2008), 281–297.

[10] Hennessy, K., Lyons, N., Loring, S., Arnold, C., Joe, M., Elias, A., & Pokiak, J. The Inuvialuit Living History Project: Digital Return as the Forging of Relationships Between Institutions, People, and Data. *Museum Anthropology Review*, *7*(1–2), (2013), 44–73.

[11] Hennessy, K., Wallace, R., Jakobsen, N., & Arnold, C. Virtual Repatriation and the Application Programming Interface: From the Smithsonian Institution's MacFarlane Collection to Inuvialuit Living History. In Proc. *Museums and the Web 2012.* San Diego: Archives & Museum Informatics (2012).

[12] Hill, T., & Nicks, T. Task Force Report on Museums and First Peoples. Canadian Museums Association and the Assembly of First Nations. Ottawa, Ontario: Canadian Museums Association, 1992.

[13] Hollinger, R. E., John Jr, E., Jacobs, H., Moran-Collins, L., Thome, C., Zastrow, J., … Rossi, V. Tlingit-Smithsonian Collaborations with 3D Digitization of Cultural Objects. *Museum Anthropology Review*, *7*(1–2), (2013), 201–253.

[14] Hornecker, E., Marshall, P., & Rogers, Y. (2007). From entry to access: how shareability comes about. In *Proc. DPPI 2007*, ACM Press (2007), 328.

[15] Iversen, O. S., Halskov, K., & Leong, T. W. Values-led participatory design. *CoDesign*, *8*(2–3), (2012), 87–103.

[16] Jun, S., & Lee, H.-K. Dialogue and carnival: understanding visitors' engagement in design museums. *Digital Creativity*, *25*(3), (2014), 247–254.

[17] Kurin, R. Safeguarding Intangible Cultural Heritage in the 2003 UNESCO Convention: a critical appraisal. *Museum International*, *56*(1–2), (2004), 66–77.

[18] Leopold, R. Articulating Culturally Sensitive Knowledge Online: A Cherokee Case Study. *Museum Anthropology Review*, *7*(1–2), (2013), 85–104.

[19] Muntean, R., Hennessy, K., Antle, A., Rowley, S., Wilson, J., & Matkin, B. ʔeləw̓k̓ʷ – Belongings: Tangible Interactions with Intangible Heritage. *Journal of Science and Technology of the Arts*, *7*(2), (2015), 59.

[20] Muntean, R., Hennessy, K., Antle, A., Rowley, S., Wilson, J., Matkin, B., ... Wakkary, R. ʔeləw̓k̓ʷ – Belongings: A Tangible Interface for Intangible Cultural Heritage. In *Proc. EVA 2015* (2015).

[21] Phillips, R. B. *Museum pieces: Toward the Indigenization of Canadian museums* (Vol. 7). McGill-Queen's Press-MQUP, 2011.

[22] Pratt, M. L. Imperial Eyes: Travel Writing and Transculturation. New York: Routledge, 1992.

[23] Roy, S. *These mysterious people: shaping history and archaeology in a Northwest Coast community* (Vol. 64). McGill-Queen's Press-MQUP, 2010.

[24] Shankar, G., & Hooee, C. Zuni Cultural Heritage Materials in the American Folklife Center: The Potential of Return. *Museum Anthropology Review*, *7*(1–2), (2013), 74–84.

[25] Simon, N. (2010). *The Participatory Museum*. Museum 2.0., Santa Cruz, California, 2010.

[26] Smith, R. C., & Iversen, O. S. Participatory heritage innovation: designing dialogic sites of engagement. *Digital Creativity*, *25*(3), (2014), 255–268.

[27] Wilson, J. P. H. (2015). *sq̓əq̓ip : gathered together*. University of British Columbia.

Towards Interaction Around Unmodified Camera-equipped Mobile Devices

Jens Grubert
Coburg University
jg@jensgrubert.de

Eyal Ofek
Michel Pahud
Microsoft Research
eyalofek@microsoft.com
mpahud@microsoft.com

Matthias Kranz
University of Passau
matthias.kranz@uni-passau.de

Dieter Schmalstieg
Graz University of Technology
schmalstieg@tugraz.at

Abstract
Around-device interaction promises to extend the input space of mobile and wearable devices beyond the common but restricted touchscreen. So far, most around-device interaction approaches rely on instrumenting the device or the environment with additional sensors. We believe, that the full potential of ordinary cameras, specifically user-facing cameras, which are integrated in most mobile devices today, are not used to their full potential, yet. We To this end, we present a novel approach for extending the input space around unmodified mobile devices using built-in front-facing cameras of unmodified handheld devices. Our approach estimates hand poses and gestures through reflections in sunglasses, ski goggles or visors. Thereby, GlassHands creates an enlarged input space, rivaling input reach on large touch displays. We discuss the idea, its limitations and future work.

Author Keywords
around-device interaction; cross-device interaction; mobile interaction; wearable interaction; glasses; reflection; corneal imaging

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User Interfaces - Graphical user interfaces

Introduction

Handheld and wearable touch displays allow us to interact in a multitude of mobile contexts. However, shrinking device sizes, aiming at increased mobility [4], often sacrifice the interactive surface area. If devices shrink, while fingers stay the same, interaction may become inefficient. Hence, there is a need for compensating for the lack of physical interaction area.

One option is to decouple input and output area of interactive displays, using sensors to increase the input area around devices [8], extending it to near-by surfaces or to mid-air. Numerous research has sparked in the area of around-device interaction. So far, most research focused on equipping either mobiles [1], the environment [10] or the user [3, 5] with additional sensors. However, deployment of such hardware modifications is hard. Market size considerations discourage application developers, which limits technology acceptance in the real-world [2].

We envision a future in which *unmodified* mobile and wearable devices that are equipped with standard front-facing cameras allow for ample movements, including the environment around and to the sides of the device, without the need for equipping the device or the environment with additional sensing hardware.

While there is a large body of work on around- and cross-device interaction most of them have the need for physically modifiying the device or equipping the environment with additional hardware [6]. Song et. al [11] enabled in-air gestures using the front and back facing cameras of unmodified mobile devices. However, their interaction space is limited to the field-of-view of the cameras, constraining the interaction space to two narrow cones in front and behind a device. Much of the interaction space around the mobile device, such as the areas to the sides of the device are not observed by these cameras. The closest work to ours is Surround See by Yang et al. [12]. They modified the front-facing camera of a mobile phone with an omnidirectional lens, extending its field of view to $360°$ horizontally. They showcased different application areas, including peripheral environment, object and activity detection, including hand gestures and pointing, but did not comment on the recognition accuracy. Their approach, just like ours, supports a large interactive space around the device. The need to add non-standard hardware to the phone limits the deployment of this technology. Furthermore, the $360°$ lenses used by Yang et al. increase the size of the device thickness, making it hard to access and store a mobile device. In contrast, our approach only requires access to common and widely available apparels, which can result in a software only deployment to enable around-device interaction on millions of existing mobile devices.

In contrast to previous work, we propose to enrich the sensing capabilities of unmodified mobiles by everyday common apparels such as sunglasses or common reflective visors.

Our interactive system, called *GlassHands*, utilizes reflective glasses or visors to extend the field-of-view of front-facing cameras built into mobile devices, mimicking effects of catadioptric panorama cameras [7]. Other reflective surfaces, such as ski goggles, diving masks or helmet visors, may enable gesture interaction with phones as well. This is of interest, when fine interaction with small screen is dangerous or impossible, for example, when the hands are covered with gloves. In fact, for some scenarios, such as skiing, reflective visors are so common, that a software-only deployment may also be economically feasible.

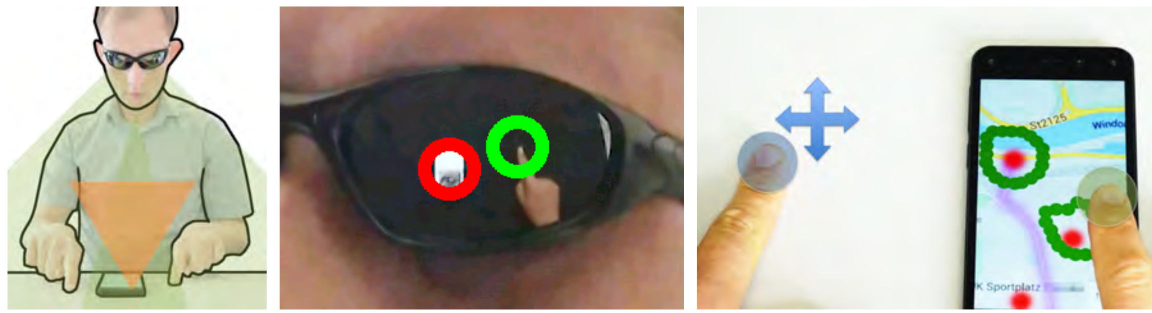

Figure 1: GlassHands extends the input space around mobile devices. Left: The narrow field-of-view of front facing cameras (orange) is extended through sunglasses reflections (green). Middle: Detected mobile phone (red) and finger tip (green) in glass area. Right: Users can continuously pan outside the display while simultaneously tracing over items on the device screen.

Around-Device Interaction using Reflective Glasses

Off-phone interaction requires sensors that can observe the interaction, in our case, touch events on a surface and in-air events around the phone. However, the unmodified phone does not have any such sensing ability. GlassHands mimics a virtual external point-of-view (POV) that captures large areas of the workspace, by using the existing camera to observe a reflection from surface that lies in front of the camera. The reflection will contain the phone itself, the surface around the phone, and the user's hands.

Such a reflectors may be part of the environment, mounted on a ceiling or above a workstation. In this work, we look at *wearable* reflectors such as glasses or visors. Being worn in front of the user eyes, these reflectors are naturally positioned to face the phone. In many useful scenarios where normal touch interaction is difficult, for example, when the user has to wear gloves, or dangerous, such as skiing, bike riding, diving, manufacturing and more, such eye-wear is already common and can be leveraged.

To use the reflection image as an input modality, we need to be able to detect the reflected image in the camera image and extract the relative location of objects that need to be sensed. Specifically, we must detect the position of the user's hands next to the phone and, finally, map this information to world coordinates.

The components of the system encompass detection of the user head and the area of the user's glasses where the reflection is visible, detection of the phone and the user hands in the reflection, and translation of all these locations to metric world coordinates around the phone. Finally, the implementation of a touch interface requires the ability to detect when the user's hands touch the surface.

We have implemented the system on an Amazon fire phone using OpenCV for image processing and HTML5 and JavaScript for application development. We note, that while our system was implemented on an Amazon fire phone, it can be employed on other commodity smartphones as well. The system workflow is summarized in Figure 2.

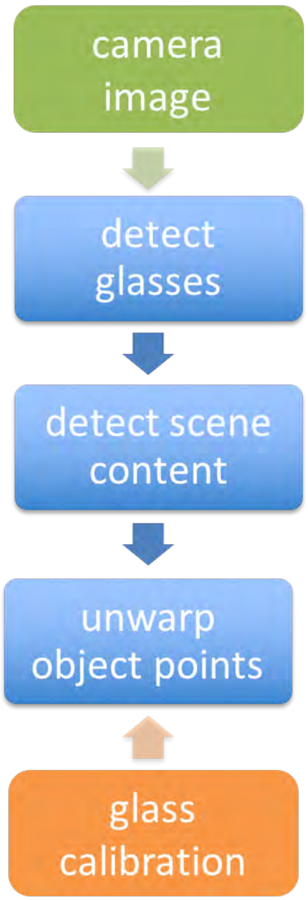

Figure 2: GlassHands processes image data (green) to determine the glasses region and scene content. Calibration data (orange) is used to transform phone and hand coordinates from the camera's image space to the display space of the phone.

Applications

We demonstrate the potential of our concept by implementing various application prototypes, which are described next.

Map Navigation

Touch-based map navigation on mobiles is limited by the small screen space: Panning to distant objects or zooming large distances typically involves repeated drag and pinch gestures. Wearing gloves, for example while skying or biking, may prevent using touch to interact with the phone. Using GlassHands, one can pan and zoom utilizing the surface, or the air around the device (see Figure 3) or simultaneously pan and trace over an item of interest (see Figure 1, right).

Moreover, navigating a map by outside device gestures avoids occlusion by the fingers, which is an advantage for small displays. Touching solely inside or outside the display causes the map to be paned and zoomed normally.

Task Switching

Switching tasks on mobiles involves typically at least two touches and an additional pan action, searching for the requested application among a list of prior used applications. Furthermore, if task switching includes cutting and pasting items via a pop-up menu, at least six consecutive touch actions are needed.

We allow browsing application using a linear ribbon metaphor (see Figure 3) and enable fast cutting and pasting: The user selects an item to be moved, and holds it. Next, the user switches to the target application, either directly through tapping on locations as mentioned above or by panning the ribbon with a finger drag outside the phone. Upon reaching the target application, the user releases the item to paste it into the target. Prior to releasing, the user can place the item at the desired location within the target application by dragging it on the phone's display. This operation can also be done in mid-air, where the holding hand thumb is used for item selection.

Figure 3: Top: Panning and zooming a map requires fewer repetitive gestures, compared to touch-screen only interaction and avoids screen occlusions. Bottom: A virtual ribbon allows browsing open applications, by panning outside the screen.

Mid-Air Music Player

Working with touch screens of mobile devices while wearing gloves, such as while skiing, viewing a map on a motorbike (at a traffic stop) or answering a phone while using work gloves, can be cumbersome. Users have to take off their gloves to operate common capacitive touch screens, e.g., when browsing through a music collection or when unlocking the screen. While touch screen capable gloves exist, they are definitely rarer than ski goggles. Furthermore, touching mobile screens with gloves can lead to an amplified fat finger problem. We have implemented a music player application that allows browsing music collections using mid-air gestures. The user can initiate a gesture by holding the hand next to the phone. Then, hand movements to either side of the phone are mapped to a scrolling list, as can be seen in Figure 4. The large available interaction space allows fine accuracy of selection. The same application can also be used with hand gestures on surfaces.

Discussion and Conclusion

We demonstrated that interaction at the periphery of a mobile device is feasible given a front-facing RGB camera and a user wearing sunglasses, or any other reflector. In contrast to previous work [11], we significantly broaden the interaction space around an unmodified mobile device. In contrast to Surround-See [12], we enable large interaction space, without the need for special hardware, just everyday common apparel, and with keeping the phone thin and small. Removing the need for special hardware enable us to easily deploy this solution, such as through a store application.

The proposed system has several limitations, some which may be addressed in future work:

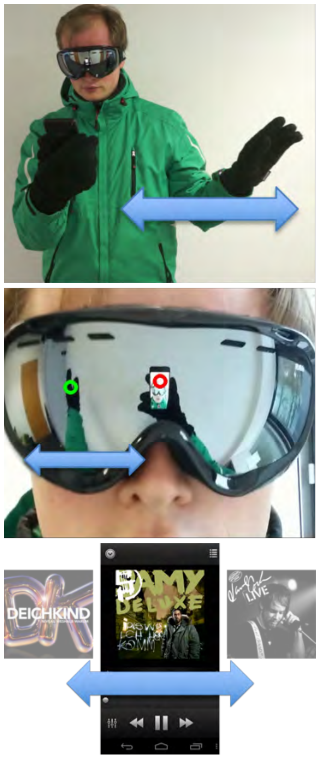

Figure 4: Top: Mid-air sliding gesture. GlassHands support large hand gestures, that enables continues scrolling or selection from a large list of options. Middle: Close-up with detected phone (red circle) and glove (green circle). Bottom: The album selection of a music player is operated with left and right sliding gestures.

An inherent limitation of our approach is that users have to wear reflective glasses, which have to be visible in the field of view of the front camera (typically 70-80 degrees field of view). Dark glasses allows the camera to view a clear reflection without the view of the user eyes behind them. If the user face is relatively dark, such as in the case of a desk lamp illuminating the table, while the user's head stays in the dark, then regular glasses can be used just as well.

In general, the usage of dark glasses in low-light environments, like indoor offices, might not be appropriate, both due to perceptual and social reasons. However, there are many situations, both indoors and outdoors, where wearing reflective eye-wear is common: workers wearing safety goggles, skiers, divers, motorcyclists and so on.

On-surface interactions as described above enable comfortable work, where the users hands are supported by the surface, and the tap of a finger on the surface can be used to detect touch. However, the technology is not limited to surface-based interaction. The same 2D interaction may be used in mid-air around the phone with ample gestures. Such in-air interaction with Glasshands could help skiers operate their mobiles without the need to take off their gloves.

Also, GlassHands could be used by bikers, who have mounted their phone on the handle bar. Sliding the hands along the handle bar, relatively far from the phone, could be utilized to steer on-screen applications (see Figure 4, right). A possible application may sense the direction of a pointed hand, reflected in the helmet visor, and use the phone GPS and orientation to announce the street number of the house the biker is pointing at. Divers could use GlassHands to interact with their phone, stored inside a watertight casing. In a similar fashion, workers who wear protective glasses and gloves may interact using large gestures around the phones.

Measuring the hand positioning accuracy shows that a minimal distance of at least 5 cm is needed to reliably separate two individual touch down events on the surface, when a single frame is used for measurement. It is possible to increase the measurement accuracy, for example, using temporal filtering.

Furthermore, we deliberately employed simple and efficient computer vision techniques throughout our pipeline to demonstrate the feasibility of the GlassHands concept. These simple algorithms can not cope well with complex environments, typically found in real-world situations. For practical implementations in commercial applications, more robust algorithms should be used (see also below).

At the camera resolution we used (0.9 megapixel), the image of each lens is about 130 by 170 pixels, which limits the maximum accuracy of the hands location to about 0.5 cm. Better front cameras are already available and should improve the quality of detection and location estimation.

In the future, we want to support a wider variety of glasses models with different reflection and curvature properties. Moreover, the use of reflections directly of the user's eye using corneal imaging [9] could overcome the need for eyewear. Figure 5 shows the reflection of the user eye, as captured by a 6 megapixel DSLR placed on a table at the same distance from the user as the phone. Hands are detected using a standard hand tracking approach in the cornea of the user. Currently available mobile phones do not have such high resolution front cameras, but future models may support them, making the presented technique fit more usage scenarios and remove the need of wearing reflective glasses.

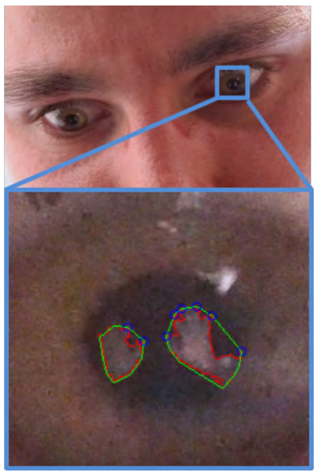

Figure 5: Top: 6 megapixel image of a users face area. Bottom: Hands are detected in the cropped corneal area (0.16 megapixel).

Furthermore, by using both lenses of the glasses, one could model two catadioptric camera systems and apply stereo techniques to recover depth values. We plan to investigate how to estimate such camera models on the fly. Depending on the quality of the stereo reconstruction, usage may range from determining the height of the user's hands above the surface to possible replacement of body-worn depth sensors.

GlassHands allows to sense an ample area in a phone's vicinity. There are options to recognize objects in space and react accordingly. The interaction may involve everyday objects, such as toys on a table, a board-game, or ingredients on a kitchen counter. As a next steps, our approach could be extended to ad-hoc spatial registration of multi-display environments without the need for external cameras [10].

We believe that new sensing capabilities for phones will help spreading spatially aware applications. In many cases, the development of such applications is hampered by the limited distribution of required hardware. Hardware manufacturers, on the other hand, may hesitate to include new technology without proven value for applications [2]. We hope that an approach such as GlassHands may break this circle by enabling applications aimed at a specific scenario, such as Skiing, to be commercially successful using existing hardware. Ultimately, this may encourage the development of new dedicated sensors integrated in consumer phones.

REFERENCES

1. Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight: Multi-"Touch" Interaction Around Small Devices. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST '08)*. ACM, New York, NY, USA, 201–204. DOI: http://dx.doi.org/10.1145/1449715.1449746

2. Bill Buxton. 2008. The long nose of innovation. *Insight* 11 (2008), 27.

3. Liwei Chan, Chi-Hao Hsieh, Yi-Ling Chen, Shuo Yang, Da-Yuan Huang, Rong-Hao Liang, and Bing-Yu Chen. 2015. Cyclops: Wearable and Single-Piece Full-Body Gesture Input Devices. In *Proc. CHI '15 (CHI '15)*. ACM, New York, NY, USA, 3001–3009. DOI: http://dx.doi.org/10.1145/2702123.2702464

4. F. Gemperle, C. Kasabach, J. Stivoric, M. Bauer, and R. Martin. 1998. Design for wearability. In *Proc. ISCW'98*. 116–122. DOI: http://dx.doi.org/10.1109/ISWC.1998.729537

5. Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. MultiFi: Multi Fidelity Interaction with Displays On and Around the Body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3933–3942. DOI: http://dx.doi.org/10.1145/2702123.2702331

6. Jens Grubert, Matthias Kranz, and Aaron Quigley. 2016a. Challenges in mobile multi-device ecosystems. *mUX: The Journal of Mobile User Experience* 5, 1 (2016), 1–22. DOI: http://dx.doi.org/10.1186/s13678-016-0007-y

7. Jens Grubert, Eyal Ofek, Michel Pahud, Matthias Kranz, and Dieter Schmalstieg. 2016b. GlassHands: Interaction Around Unmodified Mobile Devices Using Sunglasses. In *Proceedings of the ACM 2016 ACM International Conference on Interactive Surfaces and Spaces (ISS '16)*. ACM, New York, NY, USA, to appear.

8. Chris Harrison. 2010. Appropriated Interaction Surfaces. *Computer* 43, 6 (June 2010), 86–89. DOI: http://dx.doi.org/10.1109/MC.2010.158

9. Christian Nitschke, Atsushi Nakazawa, and Haruo Takemura. 2013. Corneal imaging revisited: An overview of corneal reflection analysis and applications. *Information and Media Technologies* 8, 2 (2013), 389–406. http://dx.doi.org/10.2197/ipsjtcva.5.1

10. Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. ACM, New York, NY, USA, 45–54. DOI: http://dx.doi.org/10.1145/2669485.2669500

11. Jie Song, Gábor Sörös, Fabrizio Pece, Sean Ryan Fanello, Shahram Izadi, Cem Keskin, and Otmar Hilliges. 2014. In-air Gestures Around Unmodified Mobile Devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 319–329. DOI: http://dx.doi.org/10.1145/2642918.2647373

12. Xing-Dong Yang, Khalad Hasan, Neil Bruce, and Pourang Irani. 2013. Surround-see: Enabling Peripheral Vision on Smartphones During Active Use. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 291–300. DOI: http://dx.doi.org/10.1145/2501988.2502049

Just Scratching the Surface, the Long Road to Effective Cross-Display Interaction

Narges Mahyar
UC San Diego
San Diego, CA 92122
USA
nmahyar@ucsd.edu

Kellogg S. Booth
University of British Columbia
Vancouver, BC V6T 1Z4
Canada
ksbooth@cs.ubc.ca

Cynthia L. Girling
University of British Columbia
Vancouver, BC V6T 1Z4
Canada
cgirling@sala.ubc.ca

Ronald W. Kellett
University of British Columbia
Vancouver, BC V6T 1Z4
Canada
rkellett@sala.ubc.ca

Abstract
There are five issues that face designers of systems that support cross-surface interactions "in the wild." These present unique challenges to successfully deploying multiple displays that fully exploit surface technologies and the rich interactions they afford: (1) the form factor of a display often determines its appropriate role in a multi-surface environment; (2) placement rules, replication, and presentation format for content that is shared across surfaces can have complex semantics that need careful design to be effective; (3) the physical and logical topology of linked surfaces impacts how cross-surface interaction will be controlled; (4) the rapid convergence of computer graphics, computer vision, and haptic input and output are opening up vast new possibilities that were only imaginable a few years ago; and (5) the desire to make these new technologies accessible to a widely diverse set of stakeholders makes all of the issues that much more challenging. We illustrate our discussion through examples drawn from our own work supporting collaborative urban design for sustainable cities.

Author Keywords
Charrette; collaboration; content sharing; form factor; interaction design; interdisciplinary; public engagement; stakeholder; surface; sustainability; urban design; user-centered design; visualization.

Figure 1: UD Co-Spaces: a multi-display collaborative environment where urban design stakeholders can simultaneously interact with multiple surfaces.

Figure 2: UD Co-Spaces supports 2D maps on a tabletop coordinated with 3D views on the wall

Figure 3: Google Earth is used to embed a proposed sustainable design into an existing urban context.

ACM Classification Keywords

H.5.2-3. [Information Interfaces and Presentation]: User Interfaces, Group and Organization Interfaces

Introduction

As the title of this position paper suggests, we believe that despite multiple decades of research and experience with surface computing, we have only just begun to understand what can be accomplished using multiple surfaces that are used in concert with each other. The particular issues that we focus on are those that arise from cross-display interaction. This does not mean that there are not equally important challenges for single-surface interaction. There are. But our interest is in identifying areas in which substantial progress can be made in the next few years.

The opportunities provided by computing environments with multiple display surfaces bring with them a number of challenges. Weiser first elaborated a vision of computational agents embedded in everyday settings that is now variously known as ubiquitous computing, pervasive computing, and smart environments [11]. More than a decade ago, researchers at Stanford developed the iRoom to demonstrate how a heterogeneous collection of displays could inter-operate to support seamless interactions across the displays [8]. Since then, there has been much research activity exploring collaboration on single and multiple tabletop displays for a variety of application domains.

Our own work has focused on urban design where we have studied how a multi-display environment centered on an interactive multi-user tabletop display that we call UD Co-Spaces (Fig. 1) can support the public charrette process that is commonly used by planners [2, 5].

Our main goals were: (a) engaging different groups of stakeholders to actively interact with design options, (b) fostering collaboration and co-creation of urban design through the use of touch-based interaction and surface technology, and (c) enabling stakeholders to understand consequences of their choices using simple visual encodings that connect sustainability indicators to urban form.

There has already been a lot of interest in using interactive surfaces in the context of urban planning. The archetype tabletop system for urban design was URP [9], developed by Ishii and his colleagues at the MIT Media Lab and subsequently extended to incorporate aspects of mixed reality [1, 4]. Later, Wagner et al. [10] developed a mixed-reality system for urban planning using both tabletop and wall displays with physical objects on the table. The physical objects are only used as tokens instead of being actual scale models of buildings, as in Ishii's Luminous Table [4]. These applications are more like presentation tools for previously developed urban designs, with little support for actual interactive design activities such as those that take place in brainstorming sessions when civic engagement with a wide range of stakeholders is important.

More recent tools such as ETH's ValueLab [3] bring together large interactive displays and visualizations to facilitate public participation in the planning of mega-cities – what Halatsch et al. refer to as "Future Cities." ValueLab uses GIS data and visualization software to simulate design choices. Maquil et al. developed ColorTable [6], an interactive round table to facilitate communication between diverse stakeholders. ColorTable utilises a tangible user interface and tokens placed on the table to represent design elements, such as buildings or streets. It uses a physical map augmented with 2D images and 3D objects to lend the design area more realism. A perspective view of the design is provided by a projected image on a wall display.

While these early research investigations are similar to ours

Figure 4: Discussions among stakeholders naturally flow back and forth between the multiple surfaces.

Figure 5: Sustainability indicators provide feedback to stakeholders through infographics-style widgets.

Figure 6: Indicators on an auxiliary wall display help stakeholders track progress toward their sustainabilty goals.

in the sense that they utilised multiple displays in conjunction with collaborative tabletops for urban planing, our efforts focused on providing familiar visualization and interaction techniques, accessibility, and fast and early feedback to encourage broad-based public engagement. Our assumption that ubiquitous experience with personal computers, hand-held smart phones, and tablets has created a level of sophistication in the general public that implies that gestures such as zoom and pinch will be familiar to many people was definitely borne out in our evaluations of this approach [5].

In this position paper we reflect on our experience to date and we share some of the insights that we gathered through many years of observing different iteration of UD Co-Spaces in use with different broad range of users (singles, families, youth, older adults) with a range of computer savvy from novices to experts. We believe that the insights we gained are worth considering when building future multi-display systems even for different application domains and collaborative tasks. We have structured the rest of the paper around the discussion of five key issues that we would like to bring to the attention of the cross-surface interactions community.

1. Form Factors Matter – A Lot!

The size, aspect ratio, orientation, and location (relative to viewers) of a display implies certain affordances that are different for each combination of these factors. Examples are text, which needs to be readable and thus a vertical orientation is preferred), and maps (which traditionally can be used from multiple viewpoints and thus are appropriate for horizontal orientations. Size matters at a number of scales, one being when the size is larger than a person can reach across (for a horizontal display) or reach up to (for a vertical display).

While other researchers have also investigated the importance of the spatial parameters of displays [7], in our studies with UD Co-Spaces we learned how different form factors can affect users' experience, interactions and engagement. For example, the biggest problem with the initial prototype for UD Co-Spaces was limited screen real estate. Display size was not enough for robust design exercises (*size matters*). In the second iteration, we used a bigger tabletop display and we added a 3D view (Fig. 2). Proposed buildings are viewed in context using Google Earth (Fig. 3) to facilitate discussion of the urban landscape (Fig. 4).

Content can dictate form factors when choosing a surface. The natural up-down orientation of the 3D view of a landscape and of the indicator "widgets" (Fig. 5) we designed was more appropriate for a vertical display (*orientation matters*). After our initial design we decided to move the indicators from the tabletop to the wall display (Fig. 6) to free up valuable space on the tabletop, but later we learned that it created a new display real estate problem: indicators took space on the wall that was better be used for 3D. We learned that the location of indicators affected engagement: users were so much engaged with the tabletop that they neglected the indicator metrics on the wall (*location relative to the viewer matters*). We addressed these issues in the third iteration of our system by adding personal hand-held devices to control the 3D view on the wall and to view the indicator widgets (Fig. 7), thus freeing up the wall display for 3D while also encouraging more interaction with the sustainability indicators.

2. Content Sharing and Collaboration Patterns

A similarly range of concerns relates to how display content is shared across or on surfaces, some correlated with size (small surfaces are difficult to share, very large surfaces are easy to share), but in some cases the issues are more

Figure 7: UD Co-Spaces integrates tabletop, wall, and personal hand-held surfaces.

Figure 8: Hand-held surfaces provide personal “scratch space” to set targets for indicators.

Figure 9: UD Co-Spaces uses a message-based protocol to link loosely coupled applications – perhaps running on multiple computers that are connected across a network.

about privacy, personal ‘working space’ or who has the right to determine the content of the display (or a portion of the display).

In our work, we provided personal working spaces on iPads (Fig. 8), where users could access, customize, interact and learn about the sustainability indicators. This not only allowed some personal explorations, but also allowed more parallelism where group worked on a loose collaboration style, or formed sub-groups that allowed them to work on the task in parallel. The same notion was true about the 3D view. We observed that some group members became specialized with one of the display content and interaction types, which enabled groups to go forward with collective effort. In terms of sharing, we observed that some expert users projected their iPad screens on the wall to share and discuss the indicators with another group member.

3. Physical and Logical Surface Connectivity

An entirely separate set of issues is how multiple displays can be connected, either physically or logically. Work (mostly by others) about how content or workflow moves from screen to screen includes extended desktop models and the iRoom [8] where a mouse cursor moves off one display onto another in a seamless manner that depends on the topological connections between the displays vs. what we do in UD Co-Spaces where personal handheld displays control shared wall displays through a logical connection that is independent of display topology.

With multiple wall displays there is opportunity to coordinate their content in various ways. One feature we support in the UD Co-Spaces system is having two or more 3D wall displays whose viewing parameters are synchronized. One example is three large displays each located on a different wall of a rectangular room. The 3D view of the display in the middle can be controlled from the tabletop (using the widget we designed for this) or from an iPad (using our custom app). The displays on each side have the same “look-at” point (the position on the ground plane where the virtual camera is looking) but their headings (direction from the look-at point to a virtual camera) each differ by 90 degrees.

As the middle display moves through the scene the two side displays show orthogonal views. In testing this feature we found that the display on the left should show a view from the right, and the display on the right a view from the left – at first counter-intuitive but upon reflection we realized that our paradigm was that each display is looking at the exact-same point on the ground plane, so the left-is-left and right-is-right mapping to screens that is used in a flight simulator is not really appropriate for visualizing a 3D urban design.

Important issues include how two or more displays come into these relationships. In the iRoom [8], displays are static in their physical locations, so the room ‘knows’ the topological relationships, but in the newest version of the iPad controller we need to tell it which display it is controlling or that it is controlling multiple displays at the same time. The same holds when an iPad acquires ownership of an indicator widget. We have not yet developed a full framework for this, but we have thought about the interaction sequences necessary for an iPad to select a display or an indicator, similar to how a PC running Windows selects one of multiple external displays when setting up an extended desktop. There could be multiple similar indicators so it would be important to be able to identify which instance is being modified among those that are displayed publicly.

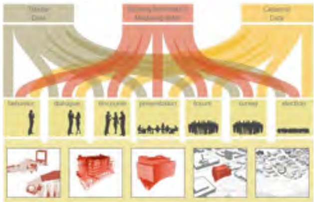

Figure 10: UD Co-Spaces is part of an on-going exploration of digital tools for sustainable urban design.

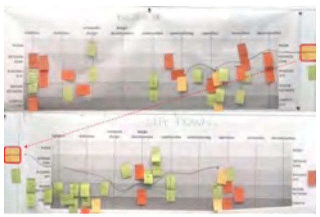

Figure 11: Existing paper-based charrettes “flow” content across surfaces as in this photo of two adjacent bulletin boards with a timeline and Post-it ® notes.

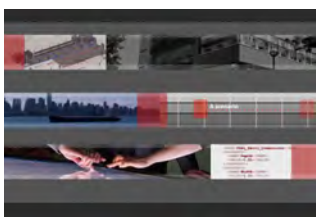

Figure 12: Urban design uses a variety of representations and visualizations within the charrette process that can be mimicked and improved using digital surfaces.

4. Rapidly Converging Technologies

A fourth issue is the convergence of computer vision and computer graphics, so that interactive surfaces in the future can be expected to be both displays and sensors that are able to see what is in front of them and use that information not just for things like gestural input but also to figure out where all of the other displays – and the people – are located within the room.

Some devices will be only one or the other (display or sensor), but more and more (like the iPad) they will have both capabilities. For example, an iPad could work much like the original 3D viewing widget did on the tabletop, but this time using vision (on the iPad) to determine the look-at point on the table and the heading and even the elevation and range by figuring out the POV parameters by matching what the iPad camera sees with what the tabletop is known to be displaying – and perhaps adding some “fiducial” information to assist the human who is trying to position the 3D camera for a wall display.

5. Diversity of Stakeholders

Matching the interaction techniques to the stakeholders is perhaps the most significant challenge when working in the wild because the range of stakeholders is large and often there is no opportunity to know who they will be until the system is actually deployed in use. The trade-off between creating sophisticated interactions to enable more personal exploration vs. keeping it simple for engaging lay people is ever present. As we noted at the outset, our hunch that stakeholders today are more savvy about touch surface technology turned out to be right. It is a pretty good bet that this will continue to be the case so that interactive surfaces for specific application domains can rely on a certain base level of expertise within many members of the public. But there will almost certainly always be a need to ensure that no segment of the stakeholders is disenfranchised by the choice of technology for a system.

Current Prototype

The most recent iteration of our system, UD Co-Spaces, uses a loosely coupled set of applications that communicate with each other through a whiteboard-style messaging system (Fig. 9).

The main application runs on a multi-touch tabletop display. The application has a database of urban forms (buildings, parks, and other structures) called *cases* that can be moved onto the background urban landscape to create a *pattern*.

The application supports abstractions that can be bound to various servers for displaying maps (Bing, Google Maps, or a static map), interpreting multi-touch gestures on a number of commercial tabletop displays (Microsoft Surface, PQ Labs, or Smart Technologies), and providing 3D visualization of the cases in a pattern in context (Cesium or Google Earth).

Handheld devices (iPads) communicate via WiFi to proxy applications that serve as forwarding agents to the messaging system that integrates services on the handheld devices with services on the tabletop. So, for example, swiping gestures on an iPad can move the point of interest in the 3D view on the wall display by sending incremental changes to a forwarding agent that then uses the messaging service to broadcast the updates to the viewpoint information to the application that uses Google Earth to render the 3D scene.

The tabletop application continuously re-calculates the values of *indicators* that serve as *metrics* for assessing the quality of an urban design pattern in terms of sustainability goals. The new values are broadcast, using the messaging

Figure 13: Integration of stakeholders' comments, notes, and calculations with geographic data is common in the charette process and is easily supported and enhanced with digital surfaces.

Figure 14: Urban design charrettes that support stakeholders in envisioning future sustainable neighborhoods who traditionally use non-interactive surfaces. UD Co-Spaces can support traditional workflows as well as offering enhanced features available with digital cross-surface surfaces interaction.

system. A separate application shows widgets with visualizations of the indicators on the wall display, or on applications running on the iPads that are notified by the forwarding agents when new indicator values are calculated.

When cases are added, removed, or moved to new locations on the tabletop, the tabletop application broadcasts these changes which are then used to update the Google Earth renderer.

Conclusion

Our work with UD Co-Spaces is part of a larger research project (Fig. 10). We have only just begun to understand what can be accomplished using multiple interactive surfaces. There is much yet to understand. This can only be achieved through sustained experimentation and testing via real applications in the wild. This can be a very challenging, cumbersome, imperfect and resource-intensive applied research model. One needs deep application domain expertise and strong collaborators, a context within which the application matters, and real-life participants willing to experiment while they are trying to accomplish something else – plus there is a need for "just in time" technical and programming expertise if the research questions we have identified (the five issues that we discussed) are to be pursued effectively.

The existing charrette approach to urban design already uses large surfaces where content flows from one surface to the next (Fig. 11), and a variety of representations ranging from highly realistic to abstract schematics are employed to understand the consequences of design choices (Fig. 12) as well as integration of textual comments, notes, and calculations with the geographic information for proposed urban designs (Fig. 13). The exciting challenge is to move from the non-interactive surfaces that are the norm (Fig. 14) to fully interactive cross-surface environments.

One area of future research could be examining how to reduce the overhead and increase the iteration speed associated with conducting research in the wild. We are still developing a robust infrastructure in which we are able to generate and "swap out" options fairly quickly to explore new ideas. It would be very useful if such an infrastructure could be independent of the core application domain content – in our case urban design – so that the issues could be more easily tackled in a variety of application domains to get a broader understanding of each of the five issues.

Acknowledgements

Funding for this work came from the Natural Sciences and Engineering Reseach Council of Canada (SSHRC), the Social Science and Humanities Research Council of Canada (NSERC), and the Graphics, Animation and New Media Network of Centres of Excellence (GRAND). Research infrastructure was provided by the Canada Foundation for Innovation (CFI), the Centre for Interactive Research on Sustainability (CIRS), and the Institute for Computing, Information, and Cognitive Systems (ICICS). We learned much from the urban design professionals and the citizens who used our systems in their public workshops in Revelstoke and Vancouver, British Columbia.

References

[1] Eran Ben-Joseph, Hiroshi Ishii, John Underkoffler, Ben Piper, and Luke Yeung. 2001. Urban simulation and the luminous planning table bridging the gap between the digital and the tangible. *J. Planning Education & Research* 21, 2 (2001), 196–203.

[2] Cynthia Girling, Ronald Kellett, and Shana Johnstone. 2006. Informing design charrettes: Tools for participation in neighbourhood-scale planning. *The Integrated*

Assessment Journal 6, 4 (2006), 109–130.

[3] J Halatsch, A Kunze, and G Schmitt. 2009. Value lab: A collaborative environment for the planning of future cities. In *Proc. 27th eCAADe*, Vol. 27. eCAADe, Brussels, 507–514.

[4] Hiroshi Ishii, John Underkoffler, Dan Chak, Ben Piper, Eran Ben-Joseph, Luke Yeung, and Zahra Kanji. 2002. Augmented Urban Planning Workbench: Overlaying Drawings, Physical Models and Digital Simulation. In *Proc. Inter. Symp. Mixed & Augmented Reality*. IEEE Computer Society, Los Alamitos, CA, 203–211.

[5] Narges Mahyar, Kelly J. Burke, Xiang (Ernest) Jialing, Siyi (Cathy) Ming, Kellogg S. Booth, Cynthia L. Girling, and Ronald W. Kellett. 2016. UD Co-Space: A Table-Centred Multi-Display Environment for Public Engagement in Urban Design Charrettes. In *Proceedings of the ACM International Conference on Interactive Surfaces and Spaces (ISS '16)*. ACM, New York, NY, USA, Article nn, 10 pages.

[6] Valérie Maquil. 2016. Towards Understanding the Design Space of Tangible User Interfaces for Collaborative Urban Planning. *Interacting with Computers* 28, 3 (2016), 332–351.

[7] Derek Reilly, Stephen Voida, Matt McKeon, Christopher Le Dantec, Jonathan Bunde-Pedersen, W Keith Edwards, Elizabeth D Mynatt, and Ali Mazalek. 2010. Space Matters: Physical-Digital and Physical-Virtual Codesign in inSpace. *IEEE Pervasive Computing* 9, 3 (2010), 0054–63.

[8] Daniel M. Russell, Norbert A. Streitz, and Terry Winograd. 2005. Building Disappearing Computers. *Commun. ACM* 48, 3 (March 2005), 42–48. DOI: http://dx.doi.org/10.1145/1047671.1047702

[9] John Underkoffler and Hiroshi Ishii. 1999. URP: A luminous-tangible workbench for urban planning and design. In *Proc. SIGCHI '99*. ACM, NY, 386–393.

[10] Ina Wagner, Maria Basile, Lisa Ehrenstrasser, Valérie Maquil, Jean-Jacques Terrin, and Mira Wagner. 2009. Supporting the Formation of Communities of Practice: Urban Planning in the MR-Tent. *Sprouts: Working Papers on Information Systems* 9 (2009), 1–11. Issue 17.

[11] Mark Weiser. 1999. The Computer for the 21st Century. *SIGMOBILE Mob. Comput. Commun. Rev.* 3, 3 (July 1999), 3–11. DOI: http://dx.doi.org/10.1145/329124.329126