Leveraging Physical Human Actions in Large Interaction Spaces

Can Liu

Leveraging Physical Human Actions
in Large Interaction Spaces

Can Liu
can.liu@lri.fr

Univ Paris-Sud & CNRS (LRI)	INRIA	Telecom ParisTech & CNRS (LTCI)
F-91405 Orsay, France	F-91405 Orsay, France	F-75013 Paris, France

ABSTRACT
Large interaction spaces such as wall-size displays allow users to interact not only with their hands, like traditional desktop environment, but also with their whole body by, e.g. walking or moving their head orientation. While this is particularly suitable for tasks where users need to navigate large amounts of data and manipulate them at the same time, we still lack a deep understanding of the advantages of large displays for such tasks. My dissertation begins with a set of studies to understand the benefits and drawbacks of a high-resolution wall-size display vs. a desktop environments. The results show strong benefits of the former due to the flexibility of "physical navigation" involving the whole body when compared with mouse input. From whole-body interaction to human-to-human interaction, my current work seeks to leverage natural human actions to collaborative contexts and to design interaction techniques that detects gestural interactions between users to support collaborative data exchange.

Author Keywords
Wall-sized displays; data manipulation; multiuser interaction.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction styles

INTRODUCTION
High resolution wall-sized displays are designed to display large amounts of data. I started my PhD by observing real users performing tasks on a ultra-high resolution wall display (5.5m×1.8m with 20480×6400 pixels). Our users are mostly professionals working with large data sets, such as neuroanatomists comparing many brain scans on the wall display to classify them or organizers of the CHI 2013 conference fine-tuning the 200+ sessions according to various criteria (Figure 1). We observed users' needs when moving items

around and grouping them in a way that was meaningful to their task.

Previous research has studied the usability of large displays for various tasks. Ball et al [1] show that larger displays promote physical navigation and improve user performance for search, navigation and pattern finding tasks. However, most previous research has addressed search, visualization and sense-making tasks [2, 1] while data manipulation tasks have been mostly ignored. When users need to navigate large amounts of data and manipulate data items at the same time, the task becomes harder to perform in traditional desktop setups because the combination of data manipulation, e.g. using pick-and-drop, and navigation, e.g. with pan-and-zoom increases the complexity of interaction. Using the wall-sized display seems easier because users can navigate by walking around and manipulate data with their hands. Users can physically "zoom in" the data by approaching the display and get an overview by stepping back. In order to analyze and understand the benefits and drawbacks of using such a wall-sized display for data manipulation and navigation tasks, my dissertation first presents a set of experiments involving both single-user and multi-user classification tasks.

Among the large body of research on gestural input, some work has focused on tracking full human body movement in collaborative settings. For example, Shadow Reaching [9] facilitates interaction at a distance by projecting a virtual shadow of the user while Sparsh [7] uses the human body as a medium to transfer data. Collocated social interaction involves not only verbal communication but also gestures and relative positions that can be detected by tracking systems to support between-user interactions, such as exchanging data. Pass-Them-Around [5] lets users share photos by throwing them towards another mobile phone. Proxemic interaction takes into account the physical distances between users and displays to improve interaction [3, 6].

Existing techniques that let users exchange digital data rely mainly on traditional user interfaces: Often times, people discuss face-to-face about some content then send emails to each other to exchange the actual data. Tabletop or public displays facilitate such digital exchanges by providing graphical representations or territories for each user. Why not support a more direct way to exchange information when users are physically present in social or collaborative environments? My dissertation proposes to augment human-to-human interactions by

Figure 1. Observation of a real task - researchers working on the CHI13 program (left) and a participant performing the controlled experiment with an abstract classification task (right)

designing new techniques that support fluid and natural cross-user interactions. To our knowledge, this approach has not received much attention in the literature.

BENEFITS OF WALL-SIZED DISPLAYS FOR NAVIGATION

While users can physically navigate a large data set in front of a wall-sized display, they can perform the same task on a computer desktop using virtual navigation techniques. We conducted a first experiment to study the trade-offs between these two types of navigation [4].

In order to compare performance of wall vs. display, we designed an abstract task that captures important aspects of real classification tasks. The task consists of moving disks to containers according to a label located in their center. It operationalizes two important factors: information density, operationalized as the text size of the labels, and task difficulty, operationalized as the number of categories. The results showed strong effects: the wall display outperformed the desktop when information density was higher and the tasks were more difficult, and the desktop performed slightly better in the opposite condition.

We performed quantitative data analysis to explain the performance gain in the wall-sized display condition. Both the wall and desktop screens displayed the exact same scene so that the exact same pixels were displayed on the wall and on the desktop when participants zoomed in to the maximum level of detail. Both physical and virtual trajectories of navigation were tracked and logged throughout the experiment. This includes the physical head movements in front of the wall and to the virtual trajectories of the view point when panning and zooming on the desktop. We computed and compared several measures to look for differences between the conditions.

The analysis of the trajectories suggested that participants operated in a larger area of the scene in front of the wall-sized display than on the desktop. This is probably due to the advantage of moving the head freely, while the monitor window constrains the view in the desktop condition. This may also explain the more qualitative differences that we observed: head trajectories follow smooth curves while virtual view-point trajectories feature sharp turns (see [4] for details.)

This experiment revealed a surprisingly strong benefit of leveraging users' natural movements for interaction. Physical navigation allows users to "zoom and pan" by walking and moving their head while performing pick-and-drop interactions with an input device. On the desktop, pan, zoom and pick-and-drop actions are performed more sequentially since the mouse is the only input device. Physical navigation therefore lets users parallelize interactions naturally.

This study raises new research questions and can inspire new interaction techniques. For example: can we design a better navigation technique when data manipulation is needed, such as a novel bi-manual interaction technique or the use of eye-tracking to improve performance on the desktop? We plan to work on this subject in the future. Another important outcome of this work is our abstract task, which can be used as a test bed for the evaluation of other techniques.

Support of Collaboration

Large interaction spaces are well suited for collaborative tasks. Previous work has studied collaborative behaviors extensively for tabletops [8]. While part of this work may apply to wall-sized displays, there are key differences due to the higher mobility of users and the fact that they are not always close to the display. We are conducting a second study with the same classification task to find out if and how a wall-sized display can support collaboration effectively.

Participants pick and drop data items with cursors on the wall that are controlled by VICON-tracked pointing devices. They perform classification tasks in pairs with three strategies: (1) *Divide and conquer:* the two participants are not allowed to talk to or help each other; (2) *Collaboration with Even Control:* participants are instructed to speak out the label they are looking for and they are encouraged to collaborate while keeping the goal of being as fast as possible. The amount of collaboration is estimated by counting the number of pick-and-drop operations done collaboratively; (3) *Collaboration with Dominant Control:* only one participant has an input device so that the other one can help but not operate. By comparing these conditions we expect to understand the gain and cost of collaboration at the interaction level on a wall display.

For a deeper understanding, we assess collaboration efficiency of wall-sized displays with a secondary factor: the *Locality* of the data layout. Locality refers to the distance between a picked item and the closest target container it can be dropped into. We expect more *local* layouts to reduce the advantage of collaboration. Locality is a relative concept that depends on the distance of the user to the wall (the layout becomes less local if the user comes closer to the display). In the previous experiment, this had a major influence on the difference in performance between wall-sized and desktop displays. We are therefore interested in finding out how locality affects collaboration. After gaining insights about the benefit and cost of collaboration on the wall, we will compare it to remote collaboration with multiple desktop.

We are currently conducting a few pilot studies. Interestingly, they seem to show lower performance in the *Collaboration with Even Control* condition when compared with *Divide and Conquer*, while *Collaboration with Dominate Control* performs better than we expected. If this result stands, it might suggest that collaboration involves a coordination cost, whereas dominant control might make collaboration more effective because users are more focused on either operating or searching. The full study will therefore hopefully help us design more effective collaborative interfaces.

## LEVERAGING NATURAL ACTIONS BETWEEN USERS

When users work on a task collaboratively in front of a wall-sized display, a user who needs to deliver information or delegate a task to another user typically needs to notify this user verbally or move virtual artifacts close to them, hence disrupting them. The other user will have to process the requested task at a later time if she is not ready to perform it right away, which is time consuming and error prone. While it is important for collaborators to be aware of the task in progress while working in parallel on different subtasks [10], the overhead of managing these coordination actions and maintaining awareness of the division of labor can be significant.

I am exploring the idea of supporting data exchange by taking advantage of the users' physical presence in the same space and their natural gestural interactions in a co-located environment. This extends the concept of leveraging body movements to multi-user contexts by augmenting human-to-human interaction with digital power.

PoPle Prototype

To instantiate the concept, I designed PoPle - an interaction technique for multiple users to exchange data and delegate tasks (Figure 2). This project is still in an early stage. The first design has been created and I have implemented a working prototype with the key functionalities.

Figure 3a) shows a sending user (left, in blue) who picks an item on the wall using her mobile device, then points to the target user (in green) and sends him the selected item. As feedback, the mobile device vibrates when a user is in the line of sight of the device. The sending user can then tap a button to transfer the item. The action is confirmed by a "drop" sound on the sender's device. The receiving user gets a different notification sound on his device.
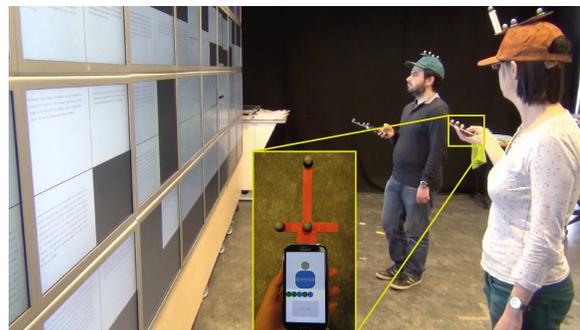
Figure 2. PoPle technique. One user(orange head) points to another user with a pointer to send him data. The pointer vibrates and shows a blue person icon (for blue head user) as well as his task load. The owner's task queue is displayed in the middle of the interface, with color-coded items to indicate the sender.

In order to minimize interruptions, items received while the receiving user is not available are buffered on their mobile device. The queue is displayed on the device and the user can pick an item by tapping it (Fig. 2). Users can choose among three levels of availability. *Immediate* mode is for users working in tight collaboration: The transfered item appears directly on the wall display and is attached to the receiving user's cursor so that he does not have to pick it up. If the receiving user already has an item attached to the cursor, the received item is added to the queue. Buffered items are automatically retrieved and attached to the cursor (one by one, in receiving order) as soon as the user releases the item attached to the cursor. *Queue* mode is intended for loose collaboration: Items are always buffered and must be manually retrieved by tapping the desired item in the queue. This makes it possible to process items in a different order than the receiving order. Finally, *Busy* mode prevents receiving items from other users.

Gestures in human-to-human communication often have subtle social meanings and implications for coordination. The current prototype highlights the items in the queue on the receiver's mobile device if the users are face-to-face when transmitting data. This way the receiver can retrieve the item easily. A common use case is when a user delegates an urgent task and talks to the other user to prioritize it. It is implemented by tracking the orientations of users to detect the face-to-face configuration.

Users can retrieve real-time information about other users' tasks in order to provide awareness of the progress of each user. Such information is important to help users monitor and equalize their workload. When pointing to a user, the sending user's mobile device shows the receiver's ID and the number of tasks on his queue (Fig. 2). Activity data related to that user can also be displayed on the wall-sized display (Fig. 3b). The system can also highlight relevant items when co-workers are close to each other (Fig. 3c).

When using raycasting to point at users, several people may be in or near the line of sight of the target user's device. To disambiguate among pointed people, we display the icons of the pointed users by increasing distance, with the selected one highlighted (Fig. 3d). The user can change the selection by moving the device closer or further away from him.
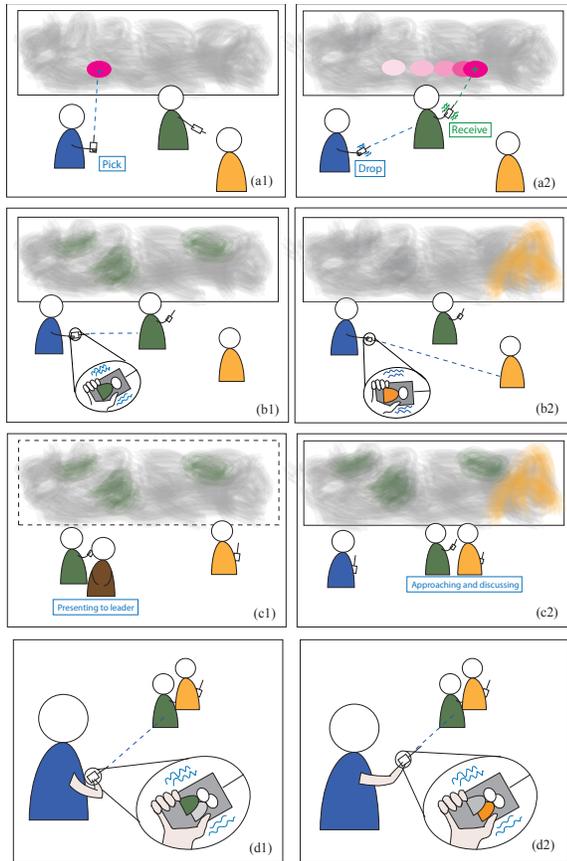
Figure 3. PoPle Techniques: Sending information (a), Highlighting related data by aiming (b) or by proximity (c), disambiguating users (d).

Design Choices

The current prototype is a starting point to explore the concept of pointing at people in collaborative settings. I plan to define a design space of such techniques that emphasizes the design and implementation choices and their effects.

For example, *directness of interaction* captures how direct, and therefore how potentially disrupting, the designation of a target user can be. It ranges from interacting with the name or icon of the person, to approaching or touching the person to sticking a physical document in their hand. The current design of PoPle is relatively direct, taking advantage of users physical presence, while providing a queuing mechanism to make it less disrupting.

Feedback and *notification* of user actions can be designed in many different ways in terms of their location, e.g. personal device of shared display, and modality, e.g. audio, tactile or visual. These choices can affect the awareness and privacy between users, as well as the availability of background information, which are known to be very important factors in collaborative tasks [10].

Finally, while the current design targets collocated collaborative interactions, the concept can be generalized to other multi-user environments. It could be used for example in an office environment or in meeting rooms to provide a more di-

rect way to share documents among co-workers. If advanced tracking technology is not available, it could use the compass available on many smartphones, or augmented reality.

CONTRIBUTION AND CONCLUSION

My dissertation will contribute to the state of the art in two main areas. First, the more fundamental research will provide a deeper understanding of the benefits of high-resolution wall-sized displays compared to desktop screens. In particular, it will show the power of leveraging human body movements for interaction. Second, the more design-oriented work will investigate the concept of augmenting human-to-human interaction with interaction techniques that detect natural social interactions to support fluid interactions among users.

I hope that this doctoral symposium will be an opportunity to discuss this work and inspire new ideas, such as: Which other human-to-human actions can be recognized and augmented with technology? How can we design systems that leverage social interactions without disrupting users?

REFERENCES

1. Ball R., North C. & Bowman D. A. Move to improve: Promoting physical navigation to increase user performance with large displays. *CHI '07*, ACM (2007).
2. Bradel L., Endert A., Koch K., Andrews C. & North C. Large high resolution displays for co-located collaborative sensemaking: Display usage and territoriality. *IJHCS 71*, 11 (2013).
3. Greenberg S., Marquardt N., Ballendat T., Diaz-Marino R. & Wang M. Proxemic interactions: The new ubicomp? *Interactions 18*, 1 (2011).
4. Liu C., Chapuis O., Beaudouin-Lafon M., Lecolinet E. & Mackay W. E. Effects of display size and navigation type on a classification task. *CHI '14*, ACM (2014).
5. Lucero A., Holopainen J. & Jokela T. Pass-them-around: Collaborative use of mobile phones for photo sharing. *CHI '11*, ACM (2011).
6. Marquardt N., Hinckley K. & Greenberg S. Cross-device interaction via micro-mobility and f-formations. *UIST '12*, ACM (2012).
7. Mistry P., Nanayakkara S. & Maes P. Sparsh: Passing data using the body as a medium. *CSCW '11*, ACM (2011).
8. Scott S. D., Carpendale M. S. T. & Inkpen K. M. Territoriality in collaborative tabletop workspaces. *CSCW '04*, ACM (2004).
9. Shoemaker G., Tang A. & Booth K. S. Shadow reaching: A new perspective on interaction for large displays. *UIST '07*, ACM (2007).
10. Yuill N. & Rogers Y. Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. *ACM ToCHI 19*, 1 (2012).