Multimodal Interaction Framework Design for Collaborative Visualization Environments

Sriram Karthik Badam*
Purdue University,
West Lafayette, IN, USA

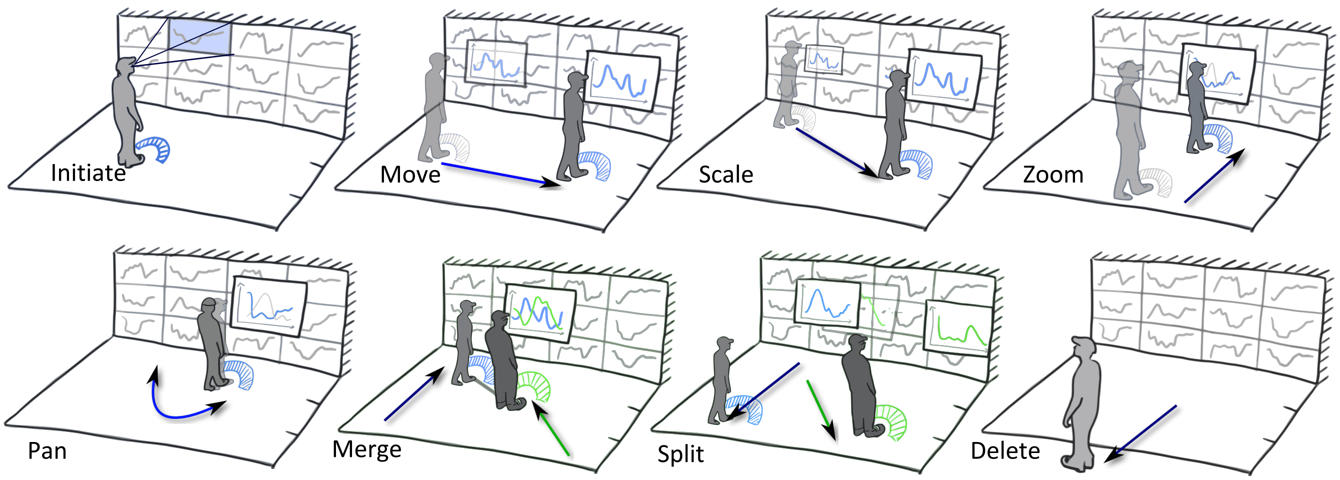

Figure 1: An example design for interaction with large display with the *proxemics* (the spatial attributes) of users and objects within the environment. In this example, the users can perform operations on a focus+context lens created from a visualization using proxemics.

Abstract

Multimodal environments are becoming increasingly common for solving various collaborative sensemaking and analytics tasks, and have gained interest as ubiquitous analytics (ubilytics) systems. My research work attempts to built interactions that span multiple devices and support multimodal input, and frameworks that can synchronize views and interaction of users at a software level in the multidevice ecosystems. I further utilize these research contributions in building a system that can bypass the known problems of various device modalities (example: limited screen space of mobile computers) by utilizing other real world objects (such as the surface of the table) to compensate for the liabilities of these devices.

1 Introduction

Ubiquitous computing has become more achievable with the surge in use of the mobile devices for computing. This has spread from the abilities to build small scale digital office environments [11] to large scale ubiquitous computing ecosystems that can connect various devices [14]. The field of visualization has started to utilize the various modalities of input and output that can enhance user experience such as large displays, tabletops, digital tablets, and smart phones, and thus ubiquitous visual analytics (ubilytics) is gaining interest [4]. Futhermore, the interaction design for these ubiquitous environments in the presence of single or multiple users has also been studied in recent works [8, 9, 12].

*e-mail: sbadam@purdue.edu

In order to fully achieve Mark Weiser's vision [13], where information processing is performed through seamless interaction with everyday objects and activites, harnesses the power of the ecosystem of digital devices around us such as large displays, mobile devices, and smart wearables is essential. Current research in the multi-device ecosystems for visualization and visual analytics focuses on building context-aware visualizations that can adapt to not only the explicit interactions performed by a user, but also the implicit information of the user's presence, mood, and preferences. For example, Jakobsen et al. [9] study how proxemics can be used to create visualizations that adapt to spatial presence of the users in terms of both the positions on the screen and also the details contained within visualizations. In contrast the work done by Chen et al. [2] in human computer interaction (HCI) that studies interaction between wearables such as smart watches and mobile devices to achieve coupling or collaboration among devices. While these attempts [5, 10] form the stepping stones for facilitating aforementioned vision of ubiquitous computing, they is still a lot of research work that needs to explored to bridge the gap between the ever improving hardware technologies and the existing interaction techniques for visualizations.

In my thesis research, I envision to build interaction techniques to both enhance interaction of the users in different collaborative coupling scenarios for analytics and sensemaking on the continuous scale of uncoupled to tightly coupled collaborative work [10]. In order to achieve this, there is a need for a framework that can allow collaboration among users and devices along with providing the ability to evaluate various interaction styles for the changing hardware modalities. This unified framework should, therefore, be platform-independent, adapt to variable display sizes, utilize all input types, and act as a middleware to provide the necessary programming abstractions to end-user developers who build multi-device ecosystems. The next few sections contain a formal definition of the problem along with existing research in solving parts of the problem, followed by my attempts to come up with an open source framework that can be used for solving the research problem by the visualization community.

Research problem and background: With the advent of wearbles such a smart watches and Google Glass [1], and increasing use of mobile devices leading to ubiquitous computing, the field of human computer interaction and visualization are trying to adapt to the improvements in the hardware technologies. For single user interaction, research done by Harrison et al. [7] promote utilization of common everyday objects such as office tables and even the human body for interaction with the limited screen space of the mobile devices. There is a need for building interactions and also adaptable frameworks that can allow platform and device-independent testing on these exemplary multimodal input devices. These frameworks should also reduce the requirement for redesigning tools completely for new device types. For example, recent work by Widgor et al. [6, 14] establishes the need for view sharing and coupling of devices at both at a framework level, where the visual interface elements in a view created for a generic device are automatically shared among devices depending on the device type, and at a interaction level where a user chooses to share specific objects (such as documents) from one device to another.

The focus of my thesis is on supporting collaborative sensemaking and analytics in multi-device ecosystems.

2 Methodology and Plan of Research

The research problem, described in the introduction, is being tackled through development of both interaction techniques and middleware frameworks for collaborative sensemaking (**step 1**), followed by integration of these attempts to build a generic system that is adaptable to technological improvements (**step 2**). The interaction design segment of this research deals with building novel interaction methods using proxemics, mobile sensor information such as accelerometer and gyroscope on smartphones, and post-WIMP interaction using depth cameras, direct-touch, and multi-touch devices. Examples of using aforementioned interaction types include using spatial position of a user to filter the contents of a visualization, and using gyroscope on mobile device as a way to rotate the view of a visualization. The frameworks for collaborative visualization handle synchronization of user interaction on one device with other devices, and also manage the display space among connected devices. Existing research work (by others) into these frameworks and interaction design include Conductor [6], Hugin [10], Lark [12], and Panelrama [14].

The second step in this research deals with the integration of working interactions into the frameworks built in step 1, that can automatically synchronize user interaction across devices, to support collaborative sensemaking and analytics. The research work in this step is under progress, and it includes building novel display and input technologies utilizing the techniques and frameworks from step 1. This step also needs usability studies of these new devices in realistic sensemaking scenarios both in terms of user experience and software delays involved in sharing interactions and building a context-aware ecosystem.

Research work that will be performed during step 2 also includes identifying liabilities of various technology and studying how each of these drawbacks are corrected by the presence of other technologies. A simple example can be explained using Google Glass. While a single Google Glass augments the real world, it cannot be adopted directly into the field of visualization, as the standard input to a Glass is in speech format (or through mobile), and only limited research work in visualization have proposed to handle interaction through speech (most of them infact only apply speech to a very specific research problem that may not be generalized). Since Glass is built for single user interaction and a simple way to adopt it to a multi-user interaction scenario is through using the connected mobile device for surrogate interaction but this may fail in many scenarios. Therefore, with the unified frameworks it is worth evaluating the ways these device liabilites can be compensated.

The research work accomplished so far in step 1 includes PolyChrome, Proxemic Lenses, and skWiki [15]. For step 2, Display Puck is the envisioned project that is currently in progress. Figure 2 shows a hierarchial listing of my research and how they merge together into the unified solution (the contribution of my dissertation).

3 Current Research

This section covers the research that was accomplished so far, and research work that is still in progress (Display Puck). The research projects undertaken as a part of step 1 described in Section 2 include PolyChrome, Proxemic Lenses, and skWiki [15]. The current research in progress for step 2 includes Display Puck. A brief description and findings from these works are described here.

3.1 PolyChrome

PolyChrome is a web application framework for sharing interactions and managing display space of various devices in collaborative settings. The framework is built entirely in JavaScript and the participating devices need only a modern web browser to utilize PolyChrome. PolyChrome uses a peer-to-peer network for sharing browser-level interactions events such as mouse clicks, keyboard presses, and direct-touch through a P2P channel, while handling persistent data such as login details, shared state, and consistency management are perfomed by a central server to reduce the load on the individual devices as most of them can be mobile devices. The framework also allows various display space configurations to handle typical multi-device environments containing large displays, tabletops, tablets, and smart phones (Figure 3).

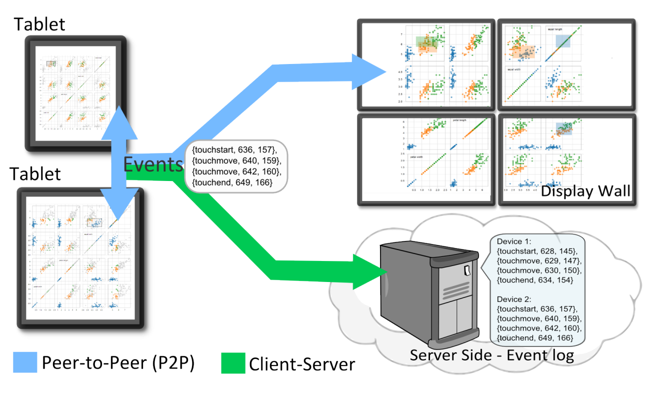

Figure 3: A web-based collaborative visualization of a scatterplot matrix created using PolyChrome. The interaction performed on the tablets are represented on the Display wall through interaction event replication. PolyChrome also allows the webpages to adapt to the multi-screen displays as shown in this mockup of a real use.

PolyChrome provides three distinct contributions to the community building collaborative visualizations: (1) It distributes legacy visualizations for the web across multiple devices and share the interactions. This can be helpful in using websites that invite social analysis of a visualization such as ManyEyes collaboratively, (2) provides the necessary interaction event and display space distribution mechanisms to create new collaborative web visualizations for multiple users and devices. This is useful for designing web visualization sites that allow all types of collaboration, including the four combinations of synchronous vs. asynchronous and co-located vs.

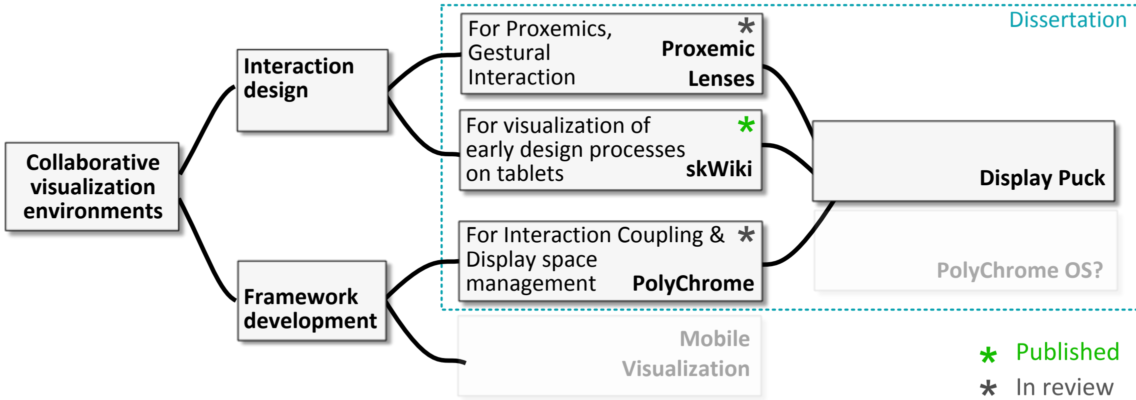

Figure 2: A hierarchial representation of my research work. Proxemic Lenses and skWiki [15] fall into interaction design described in Section 2. PolyChrome is a framework for space management and synchronizing in collaborative web visualizations. Some of the research work is still in progress and some of it has not been initiated (transparent).

distributed collaboration, (3) provides ability to store the interactions of users for replay, conflict, and consistency management.

PolyChrome has been built with the abilities to add custom interaction events like tap, slide, or even visual events such as changing the center of geographical map. This allows us to integrate PolyChrome with the interactions tested during the interaction design phase of step 1. Furthermore, preliminary testing of PolyChrome showed that the application logic of PolyChrome did **not** affect the performance of legacy visualizations to a perceivable degree.

3.2 Proxemic Lens

Research by Czerwinski et al. [3] evaluates and proves the ability of large displays to support multiple concurrent users and enhance productivity, and are therefore useful for co-located collaboration. These large displays and multi-monitor setups produce a large space to think and interact. As Harrison [7] points out mobile devices on the other hand restrict the interaction spaces, although they provide the ability of ubiquitous computing. Together these multisurface ecosystems are becoming increasingly common even in standard office environments. Novel interaction techniques [5] have started to target this very distributed nature of these ecosystems to utilize the *proxemics* information of the users and objects in their interaction models.

We created a technqiue called Proxemic lenses that can utilize the proxemics and gesture based interactions for interacting with focus+context lenses applied to visualizations, as shown in Figure 1, 4. We define the proxemics-based interaction, utilizing the spatial position, orientation, distance, and identity of the users to determine the visual attributes of a lens, as implicit interaction since its purpose is not primarily to interact with a computer system. In contrast, explicit interactions defined by gestures are actions performed with an user intended purpose. An example of implicit interaction happens when a user moves closer to a display, Jakobsen et al. [9] showcased how the system can automatically trigger a zoom command to give the user a larger view or show more details on the visualization. Based on the previous literature [9, 5], we identified eight operations on the proxemic lenses that can utilize both interaction styles: Lens Initiate, Scale, Move, Zoom, Pan, Merge, Split, and Delete. Note that each of these operations can have multiple implicit and explicit interaction mapping.

An evaluation was perfomed to observe the user experience while interacting with implicit and explicit interaction styles for each of these operations. The results of this experiment showed that the users liked to perform implicit actions (proxemics based) for Scale, Move, Pan, and Merge, while preferred explicit interactions (gesture based) for others. The results were entirely qualitative measured on the scales of preference, accuracy, intuitiveness, efficiency, collaboration ease, and physical effort.

Figure 4: Showcasing an interaction performed through proxemic lens technique. When the users get close to each other the lenses belonging to the users are automatically merged (implicit interaction).

3.3 skWiki

skWiki is created to support digital creativity in collaborative, potentially distributed, teams. It is a web application framework that allows creation of information rich ideas with sketches, text, and images. The operations perfomed for creation and updation of each idea, such as new strokes or text, are stored on a database in a structured form. As previous research work described in this proposal, skWiki is also platform independent and works on mobile devices, large displays, and even personal computers. skWiki supports mobility, collaboration among users, maintains revision history of ideas generated, and supports divergent & convergent work.

In order to facilitate interaction with various levels of operation/revision history generated by skWiki, we created a visualization called Path Viewer. In contrast to a traditional file-based content, skWiki utilizes the paths along the operation history to retrieve previous versions of an idea (similar to source control). The path viewer (Figure 5) resembles a tree-like visualization and it visualizes: (1) connections between versions of a idea/sketch, (2) shows branch and merge operations on an idea, (3) allows seamless retrieval of previous versions of a sketch. Note that the interaction with this visualization is specifically designed for mobile devices, and it utilizes the direct-touch capabilities to zoom, pan, and manupulate the content shown in the visualization. The interaction can also be performed on a personal computer or a large screen using a mouse. Furthermore, an evaluation of the entire framework showed that the users felt that the path viewer interaction was in fact helpful in retrieving versions of a sketch quickly. Although skWiki does not fully represent an attempt at collaborative sensemaking, it does showcase an application of collaborative web visualization which supports interaction with the ideas generated during a design process using mobile devices (tablets) and large displays.

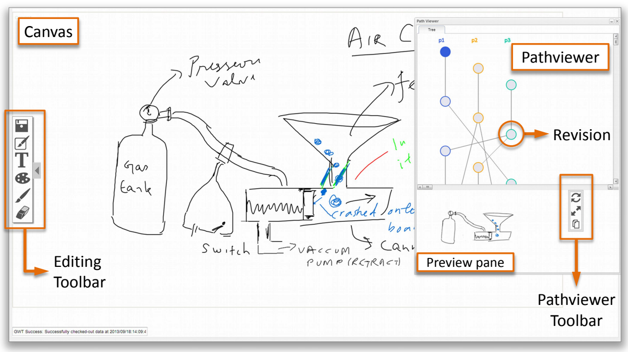

Figure 5: A screenshot of the skWiki interface with the path viewer to represent versions of sketches/ideas, and interact to retrieve, store, branch, and merge ideas.

3.4 Display Puck (Step 2 - work in progress)

Over the past decade, portable machines such as smartphones, tablets, and laptops have been progressively miniaturized, and we have reached a point where further size diminution would adversely affect the usability of these devices (as discussed by Harrison [7]). In order to enhance the user experience while working on these mobile devices, we started to utilize the physical space around us to interact with mobile devices. Display Puck, a combination of a projector and depth camera, is currently under development and it targets to solve various problems that can occur while performing analytics on a mobile device (for example, reading financial sector news and using stock market data to find smart investment choices). In combination with a mobile computer, the puck can, not only turn any real world surface into a multi-touch surface, but also allow us to collaborate with other puck users by stitching the display and input spaces. The depth camera is responsible for capturing the input from the user, and since the field of view of the depth camera is much larger than the projector it can also capture user presence. We are starting to utilize the frameworks and interaction styles created in step 1 (described in Section 2), to prototype and evaluate various interaction techniques on the puck. Its particularly interesting to see (1) how collaboration can be enhanced by building interaction for between-puck communication, (2) how devices that inherently are capable of handling input and display (projector), can enhance ubiquitous sensemaking and analytics with augmented reality (example, projecting information on a physical sheet of paper).

4 Future Work and Conclusion

In the previous section, I have described three projects (polychrome, proxemic lenses, and skWiki) that are completed (some under review) as a part of step 1 of my thesis research, i.e., interaction design, and framework development for supporting various collaborative analytics processes in multi-device ecosystems. The current plan is to utilize these for display puck (step 2), a portable computer that can convert any real world surface into touch-enabled surface that supports collaboration and augmented reality. While display puck has the potential to enhance collaboration alongside provisions for personal computing, it still deviates from the original goal of supporting any multi-device ecosystems as the puck can only support mobile computers. In future, I plan to modify the software components of display puck to work in collaboration with large displays, tabletop, and even wearables such as Google Glass.

At this point, I am looking for suggestions regarding the the organization of my research, missing components, and any loss of clarity in defining the steps I took (and plan to take). I would also like to know of any suggestions that can help me find research directions that can enhance my current work.

Acknowledgements

I would like to thank my advisor, Dr. Niklas Elmqvist, for his contінous guidance to find the path of research that I enjoy the most, and his support through my graduate studies.

References

[1] Google Glass. `http://www.google.com/glass/start/`, accessed May 2014.

[2] X. A. Chen, T. Grossman, D. J. Wigdor, and G. Fitzmaurice. Duet: Exploring joint interactions on a smart phone and a smart watch. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 2014.

[3] M. Czerwinski, G. Smith, T. Regan, B. Meyers, G. Robertson, and G. Starkweather. Toward characterizing the productivity benefits of very large displays. In *Proceedings of INTERACT*, pages 9–16, 2003.

[4] N. Elmqvist and P. Irani. Ubiquitous analytics: Interacting with big data anywhere, anytime. *IEEE Computer*, 46(4):86–89, 2013.

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

[6] P. Hamilton and D. J. Wigdor. Conductor: enabling and understanding cross-device interaction. In *Proceedings of the ACM conference on Human factors in computing systems*, 2014.

[7] C. Harrison. Appropriated interaction surfaces. *Computer*, 43(6):0086–89, 2010.

[8] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011.

[9] M. R. Jakobsen, Y. Sahlemariam Haile, S. Knudsen, and K. Hornbæk. Information visualization and proxemics: design opportunities and empirical findings. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2386–2395, 2013.

[10] W. McGrath, B. Bowman, D. McCallum, J.-D. Hincapie-Ramos, N. Elmqvist, and P. Irani. Branch-explore-merge: Facilitating real-time revision control in collaborative visual exploration. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pages 235–244, 2012.

[11] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the ACM conference on Computer graphics and interactive techniques*, 1998.

[12] M. Tobiasz, P. Isenberg, and S. Carpendale. Lark: Coordinating co-located collaboration with information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1065–1072, 2009.

[13] M. Weiser. The computer for the 21st Century. *Scientific American*, 265(3):94–104, 1991.

[14] J. Yang and D. Wigdor. Panelrama: Enabling easy specification of cross-device web applications. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 2014.

[15] Z. Zhao, S. K. Badam, S. Chandrasegaran, D. G. Park, N. L. Elmqvist, L. Kisselburgh, and K. Ramani. skwiki: A multimedia sketching system for collaborative creativity. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 2014.