

# Multi-person Spatial Interaction in a Large Immersive Display Using Smartphones as Touchpads

Gyanendra Sharma

Troy, USA

gyanendra.sharma870@gmail.com

Richard J. Radke

Troy, USA

rjradke@ecse.rpi.edu



Figure 1: Multiple users simultaneously interacting with a large screen using their smartphones and voices, coupled with spatial information about their physical locations. This panoramic image shows a 5m tall 360-degree display wall with a 44m perimeter. Please refer to the video hosted at <https://bit.ly/2GSapuM> for detailed interaction examples.

## ABSTRACT

In this paper, we present a multi-user interaction interface for a large immersive space that supports simultaneous screen interactions by combining (1) user input via personal smartphones and Bluetooth microphones, (2) spatial tracking via an overhead array of Kinect sensors, and (3) WebSocket interfaces to a webpage running on the large screen. Users are automatically, dynamically assigned personal and shared screen sub-spaces based on their tracked location with respect to the screen, and use a webpage on their personal smartphone for touchpad-type input. We report user experiments using our interaction framework that involve image selection and placement tasks, with the ultimate goal of realizing display-wall environments as viable, interactive workspaces with natural multimodal interfaces.

## ACM Classification Keywords

H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces; H.5.1 Information Interfaces and Presentation (e.g. HCI): Multimedia Information Systems

## Author Keywords

spatial intelligence; immersive spaces; interaction design; multi person interaction; smart phones

## INTRODUCTION

Designing user interactive interfaces for large-scale immersive spaces requires accommodations that go beyond conventional input mechanisms. In recent years, incorporating multi-layered modalities such as personal touchscreen devices, voice commands, and mid-air gestures have evolved as viable alternatives [32, 25, 18, 7]. Especially in projector-based displays like the one discussed here, distant interaction via smartphone-like devices plays a pivotal role [22].

Apart from the input modes of interaction, the size and scale of such spaces greatly benefit from contextualizing user locations within the space for interaction design purposes [23, 35, 17]. This is especially true for large enclosed displays such as CAVE [11, 10], CAVE2 [14], CUBE [28] and CRAIVE [30]. Representing physical user locations on such screen spaces presents considerable challenges due to spatial ambiguity compared to flat display walls.

In this paper, we present mechanisms for multiple users to simultaneously interact with a large immersive screen by incorporating three components: users' physical locations obtained from external range sensors, ubiquitous input devices such as smartphones and Bluetooth microphones, and automatic contextualization of personal vs. shared screen areas. Discrete personal interaction regions appear on two sides of a rectangular enclosed screen, where users freely move to make spatial selections and manipulate or generate relevant images. The shared screen region between the two sides can be simultaneously used by multiple users to create a desired layout based on combinations of pre-selected text with user-curated images.

Our method and overall architecture allows multiple users to interface with the large visually immersive space in a natural way. Integrating personal devices and voice along with spatial intelligence to define personal and shared interaction areas opens avenues to use the space for applications such as classroom learning, collaboration, and game play.

We designed controlled laboratory experiments with 14 participants to test the usability, intuitiveness and comfort of this multimodal, multi-user-to-large-screen interaction interface. Based on the results, we observe that the designed mechanism is easy to use and adds a degree of fun and enjoyment to users while in the space.

## BACKGROUND AND RELATED WORK

Our system is inspired by a diverse body of prior work, generally related to spatial sense-making in large immersive spaces, personal vs. shared spaces in large screens, multi-user support, and interactions using ubiquitous devices such as smartphones.

### *Spatial intelligence in immersive spaces*

Microsoft Kinects and similar 3-D sensors have been widely used for user locations or gestural interpretation in the context of various large screens [1, 36, 2] and common spaces [6]. Research has primarily been focused on developing mid-air gestures and other interaction mechanisms using methods similar to ray-casting, which require knowledge of spatial layout and users' physical locations [19]. A unique aspect of our system is the overhead Kinect array that allows many users to be simultaneously tracked and their locations to be correlated to screen coordinates and workspaces.

### *Personal vs. shared spaces*

In terms of demarcating public vs. personal spaces within large screens, Vogel and Balakrishnan [32] discussed how public displays can accommodate and transition between public and personal interaction modes based on several factors. This thread of research extends to privacy-supporting infrastructure and technologies [8, 15]. Wallace et al. [34] recently studied approaches to defining personal spaces in the context of a large touch screen display, which we cannot directly incorporate in our system but inspired our design considerations.

### *Multi-user support*

Realizing large immersive spaces as purposeful collaboration spaces through multi-user interaction support remains an active area of research [3]. Various approaches such as visualization of group interaction [33], agile team collaboration [20], along with use cases such as board meeting scenarios [16], have been proposed. The Collaborative Newspaper by Lander et al. [21] and Wordster by Luo et al. [24] showed how multiple users can interact at the same time with a large display. Doshi et al. presented a multi-user application for conference scheduling using digital "sticky notes" on a large screen [12].

### *Smartphones as interaction devices*

The limitations of conventional input devices for natural interactions with pervasive displays have led to several innovations, for example allowing ubiquitous devices such as smartphones to be used as interaction devices. Such touchscreen devices allow for greater flexibility and diversity in how interaction mechanisms with pervasive displays are materialized. Earlier concepts such as the one proposed by Ballagas et al. [5] have evolved towards more native web-based or standalone application-based interfaces. For instance, Baldauf et al. developed a web-based remote control to interact with public screens called *ATREUS* [4]. Beyond the touchscreen element of smartphones, researchers have investigated combining touch and air gestures [9], 3D interaction mechanisms [13] and using built-in flashlights for interaction [31].

## SYSTEM DESIGN

Our system was designed and implemented in a large immersive display wall environment with a 5m tall 360-degree

front-projected screen enclosing a  $12\text{m} \times 10\text{m}$  walkable area. The screen is equipped with 8  $1200 \times 1920$  resolution projectors, resulting in an effective  $1200 \times 14500$  pixel display, and contains a network of 6 overhead Kinect sensors for visual tracking of multiple participants.

### **Spatial sense-making**

Large immersive spaces have exciting potential to support simultaneous multi-user interactions. Flat 2D displays can support such functions simply by using multiple input devices with minimal consideration for physical user locations. However, to instrument large immersive environments for multi-person usage, it is necessary to demarcate personal vs. collaborative or shared sub-spaces within the context of the large screen. Contextualizing physical user locations in the space plays an important role.

To allow multiple users to interact with the screen at the same time, the large screen is subdivided into dynamic sub-spaces based on physical user locations. The existing ceiling-mounted Kinect tracking system returns the  $(x,y)$  location of each user in a coordinate system aligned to the rectangular floor space. Although users are tracked wherever they are in the space, we enabled display interactions only for users that are located within 2 meters of the screen, as shown in Figure 2. In this way, the center of the room acts as an inactive zone, where users can look around and decide on their next steps instead of actively participating at all times.

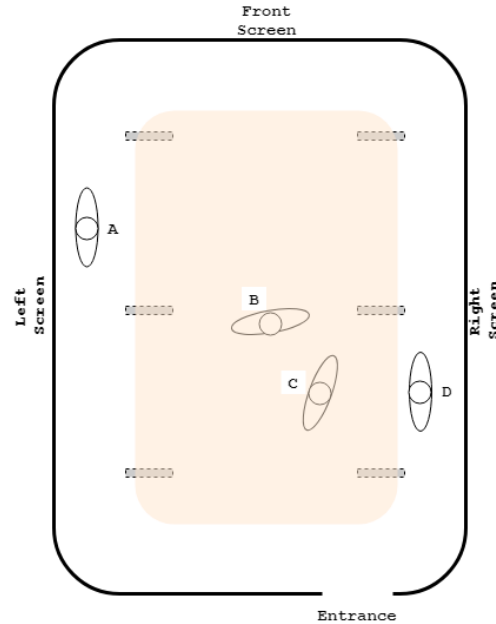


Figure 2: Users A and D are able to interact with the screen whereas users beyond 2 meters distance to the screen i.e users B and C are in the inactive region, represented in light red and thus cannot interact with the screen.

In order to make this behavior clear to the users, we carefully calibrated the floor  $(x,y)$  positions to corresponding screen locations. A key element of our design is a continuous visual

feedback mechanism shown at the bottom of the screen when a user is in range, appearing as animated circular rings, as shown in Figures 4b, 4c and 4d. This feedback serves a two-fold purpose. It makes users aware that their movements and physical locations are being automatically interpreted by the system, and it also allows them to adjust their movements to accomplish interactions with small sub-screens or columns on the large screen. Beyond the continuous feedback, we create discrete interaction spaces on the screen that change dynamically based on user locations. Thus, at a given point in time, users are able to visualize how the system continuously interprets their physical locations in real time, and also the column or sub-screen with which they are able to interact.

### Input Modes of Interaction

We experimentally explored the viability of various input methods for multiple users to interact with the large screen. These included the Leap Motion device for sensing mid-air gestures (which users found fatiguing and cumbersome to use), a fully voice-driven system (which had difficulty with some users' accents, and was discouraged by some recent studies [27, 26, 29]), and a smartwatch interface (which proved too small to easily control the large screen). Ultimately, as described below, we use each user's own smartphone as a touchpad to control the large screen, which is both familiar and intuitive to use and has an immediate personal connection.

We developed a web application that can run on any touch-screen device connected to the internet. Upon entering the environment, users showed their smartphone a QR code leading to the webpage. The webpage was designed to run as a trackpad, where familiar touch screen gestures such as tap, swipe, scroll, double tap, pinch, drag, and zoom were supported. Developing on a web platform removed the cumbersome process of users having to download and install a standalone application.

### System Architecture

The system architecture of the overall system is shown in Figure 3. It is primarily comprised of 3 components: (1) user input via smartphone and Bluetooth microphone, (2) spatial tracking via overhead Kinect sensors, and (3) the webpage running on the large immersive screen for visualization and output. All components communicate with each other in real time using the WebSocket protocol. The users' smartphone gestures are sent via WebSocket to the web application running on the large screen, as well as any voice input, which is passed through the Google speech-to-text transcription service. The user tracking system is located in a different node, which sends the  $(x,y)$  location of all users to the screen. The web application running on the large screen receives all the data, and displays dynamic feedback and visualizations accordingly.

### Overall System

Combining all the components discussed in the previous sections, we designed a multi-user spatial interaction mechanism for the large immersive space, using smartphones as input interaction devices and voice control for content generation. As shown in Figure 4, two users can walk into the space and scan

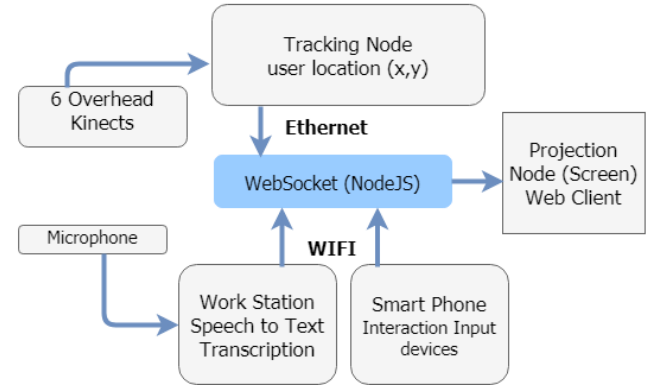


Figure 3: Overall system architecture.

a QR code located near the entrance to launch the web application on their personal devices. The two QR codes correspond to the left and right sides of the big screen. As the users move towards their respective screens and come within the defined threshold of 2 meters, the location feedback and interaction mechanisms are activated, allowing them to interact with the individual columns as they see fit. We populate each of the columns with random images from the public Flickr API. As users move around the space, the continuous spatial feedback appears at the bottom of the large screen and the column with which each user can interact is highlighted in bright red. Since the interaction column is tied to the spatial location of the user, they can be viewed as exclusive or personal to the user standing in front of it.

Phone Gestures	Screen Result
Move	Move red pointer/drag image on shared screen
Tap	Select image
Swipe (Left or Right)	Move image/s to front screen
Swipe (Up or Down)	Scroll up/down personal column
Pinch	Shrink selected image
Zoom	Enlarge selected image
Double Tap	Enlarge/shrink selected image
Long Tap	Activate/deactivate drag on shared screen
Other input	Screen Result
Move (Physical user locations)	Select different column/continuous circular visualization
Voice input ("Show me pictures of X")	Populate column with pictures of "X"

Table 1: List of phone gestures, physical user movements and voice inputs, and their corresponding screen results.

A red cursor dot that appears on the spatially selected column can be moved using the web application on the phone and acts similar to a mouse pointer. Table 1 shows the list of supported gestures and how they translate to the big screen. Images that users select on the left and right screens can be moved to the front screen, which supports a personal column for each



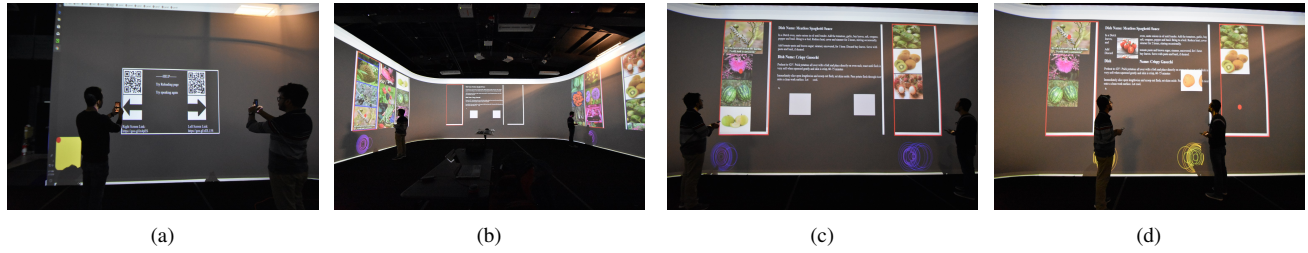


Figure 4: Multiple users during an experiment. (a) Each user scans a QR code. (b) Users work in their personal spaces using their smartphones and/or voice control. (c) Users move to the front screen to view their curated list of images. (d) Both users using the shared space to complete a full task. For a better visualization, please refer to the video in the supplemental material.

user as well as a large shared usage area. Users can move their personally curated images to the shared area on the front screen. In our particular case, we designed an application in which users can simultaneously drag their personal images around the shared screen to design a simple newspaper-article-like layout.

### USER STUDIES

We gathered 14 participants to test the usability of and gain feedback about our overall system. Only 3 participants had extensive prior experience with working in immersive display environments. We designed two experiments. The first experiment was designed to gain a quantitative understanding of how long individual users take to perform various tasks on the screen using our system. The second experiment was designed as a simple game, where two users simultaneously work using both their personal and shared screens to come up with a final correct layout. This was largely designed to understand how comfortable users felt in the space and how intuitive the felt the system to be. For this experiment, users were mostly left on their own to complete the tasks based on their understanding of how the system works.

#### Experiment 1

Individual users were directed to use only the left screen, where they were asked to complete tasks based on the prompts appearing on the large screen. There were 9 columns, each filled with random images. Screen prompts would appear randomly on any of the 9 columns asking the user to complete various tasks, one after the other. We tested all the gestures and inputs by asking the user to perform tasks shown in the second column of Table 1, except for the pinch, zoom, and double tap. Each task is completed once the user performs the correct input that corresponds to the displayed prompt.

For instance, if a user at a given point in time is in front of the 2<sup>nd</sup> column, a prompt might appear in the 9<sup>th</sup> column indicating “Select a picture from this column and move it to the front screen”. Then, the user would physically move until the system highlights the 9<sup>th</sup> column, and perform the corresponding scroll, tap, and swipe gestures. This would successfully complete the task and another prompt would appear on a different column, such as “Populate this column with pictures of dogs”, which would require a voice command. We recorded the time it took for the user to accomplish each task, including both the time it took to make spatial selections by moving between the

columns and the time it took to successfully perform phone or voice input.

#### Experiment 2

We designed this experiment to be completed in pairs. Both users had completed Experiment 1 before taking part in this experiment. Our aim was to make sure users understand all input mechanisms and are comfortable to freely use the system.

On the front screen, where the shared screen is located, we presented a simple layout with two short paragraphs of text and image placeholders. Each paragraph consisted of a heading indicating a recipe name and text below describing the ingredients and preparation. Each user was responsible for finding an appropriate image for “their” recipe. Initially, the users independently move along the left and right sides of the screen, selecting one or more images and moving them to their personal columns on the front screen. Then, they move to the front screen and select the most likely candidate image from the refined set of images and move it to the shared screen with the recipe. A screen prompt on the large screen notifies the user whether a correct image was selected (i.e., a picture of the dominant ingredient in the recipe, such as an avocado picture for a guacamole recipe). Once the correct image is moved to the shared screen, users can perform a long-tap gesture on their phone to activate dragging on the shared screen. This allows the users to simultaneously drag their answer images to an appropriate location, which is generally next to the corresponding text. A screen prompt notifies the user once the target image has been moved to the required location on the shared screen. When both users complete their tasks on the shared screen, the full task is complete.

Each user pair was presented with 6 sets of recipe “games”. 3 of the recipe pairs had the correct images already placed in one of the pre-populated columns and the users had to move around, scroll the columns, and locate the correct image. The other 3 pairs did not have the answer images in any of the columns and this required the users to generate content on their own by verbally requesting the system to populate a blank column with images of what they thought was the main ingredient in the recipe, one of which the user had to select and move to the front screen to verify.

We designed this setup to study whether users felt comfortable completing tasks based on the interaction mechanisms we designed for our display environment. We also wanted to find

out if the users, most of whom had no prior experience with these kinds of spaces, found interacting with an unconventional immersive space such as this one to be fun and intuitive. Therefore, we asked the users to fill out a NASA-TLX questionnaire along with an additional questionnaire based on a 5 point Likert scale to obtain feedback on specific spatial, gestural, and voice input mechanisms that we designed.

## RESULTS

On average, each of the 14 participants performed 27 tasks during Experiment 1, where each of the 5 tasks appeared at random. Users were required to perform at least 20 and at most 35 tasks depending on the randomness of the distributed tasks as well as their speed at completing them. All tasks were assigned equal probability of appearing, except for voice control tasks, which appeared less often, according to the design considerations discussed earlier. The average number of tasks per user was distributed as follows: spatial selection (7.35), scrolling image columns (4.93), selecting an image (6.43), moving images to the center screen (6.65), and populating with voice (2.36).

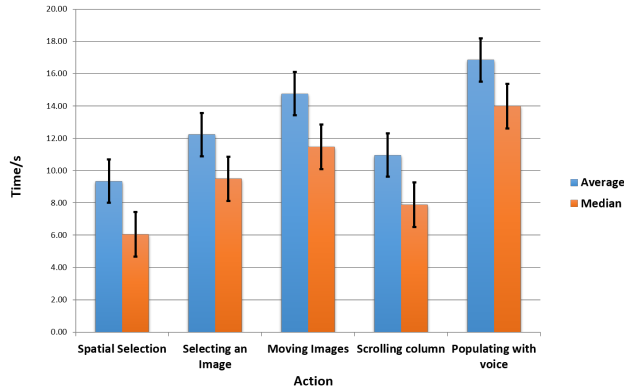


Figure 5: Average and median times for users to complete each action. All actions take longer duration than spatial selection as their completion requires spatial selection as a pre-requisite.

Even though we report both the average and median time for each of the actions, we believe that the median times for each of the tasks are more reflective of typical user performance. We observed many cases in which a certain user would take a lot of time to internalize one particular action, while completing other similar actions quickly. This varied significantly from one user to other and therefore led to some higher average values than expected. Unsurprisingly, voice input was the most time consuming action as can be seen in Figure 5.

For Experiment 2, where multiple users worked simultaneously on their personal screens and came together on the shared screen space to complete the full task, we recorded the time of completion. Since there were 3 games for the touch-only interface and 3 for the voice interface, each pair of participants played 6 games. Out of the 21 games for each type of input (7 participant pairs  $\times$  3 games per input type), participants completed 17 of each. 4 games for each input were not completed for various reasons, typically a system

crash or one of the participants taking too long to figure out the answer and giving up. The average time taken for a pair of participants to complete the touch-only based game and voice-based game were 2.31 minutes and 1.67 minutes respectively. Even though experiment 1 revealed that voice input generally takes longer, we note that for touch input the user has to physically move and search for the correct image among a wide array of choices, while for the voice input, users can quickly generate for pictures of their guessed ingredient and move one to the shared screen area.

We asked participants to fill out a NASA-TLX questionnaire after completing both experiments to investigate how comfortable and usable our overall system is, and present the results in Figure 6. We added an extra question regarding the intuitiveness of the overall system, where on the 21 point scale, a higher number indicates a higher degree of intuitiveness. Overall, participants rated their mental, physical, and temporal demand, along with effort and frustration in using the system, to be low. Performance and intuitiveness were highly rated.

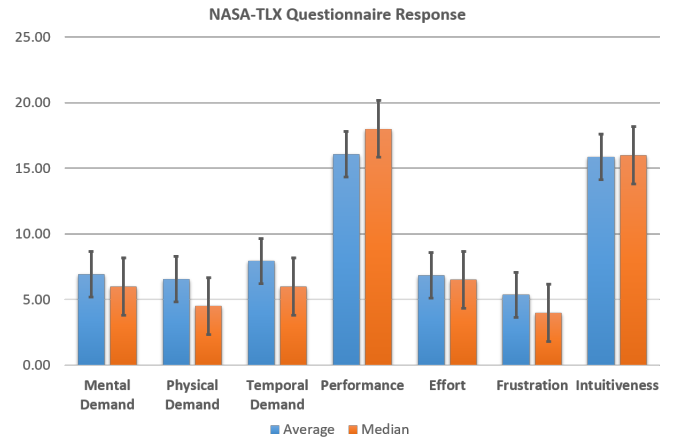


Figure 6: Average and median values of user responses to the NASA-TLX questionnaire.

In addition, users filled out another questionnaire related to how well they liked/disliked particular interaction mechanisms such as phone gestures, spatial interactions, voice input, and so on, using a 5 point Likert scale. As shown in Figure 7, median values for most of these components are rated very highly. The ratings were also high for whether the overall tasks were fun and enjoyable. Users also highly rated the user interface and other feedback on the screen, including the constant localization feedback. Among the 14 participants, 3 were previously familiar with the physical space. However, the interaction interface was completely new to them, the same as the rest of the users. We observed that the users familiar with the large immersive space performed 25% and 30% faster than the overall average for the touch and voice games respectively.

We observed that in Experiment 2, the average time for completion with voice input was less than that for the smart phone input, even though Experiment 1 revealed that voice input takes a longer time on average. This can be explained due to the time-consuming nature of search required in the phone

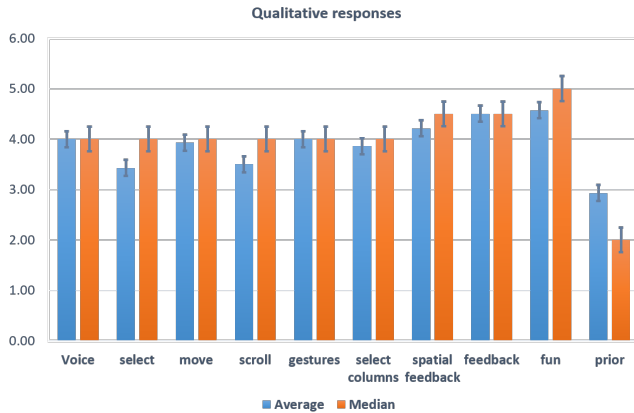


Figure 7: Average and median values of user responses to the second questionnaire on a 5 point Likert scale.

input subtask. On the other hand, for the voice input task, upon knowing the key ingredient, users were quickly able to ask for valid pictures and move them to the shared screen area.

Based on observations and post-task interviews, many users appreciated the constant spatial feedback, allowing them to understand their impact on the space. Some users appreciated the automatic demarcation of personal vs. shared within the scope of the same large screen.

We observed many issues related to the automatic speech understanding and transcription. Non-native English speakers had more difficulty populating their columns with desired input. Thus it was unsurprising that the average time for actions to be completed using voice was the largest, as shown in Figure 5. Users were divided on the usefulness and comfort of voice input; one user wished he could carry out the entire task using his voice while another was completely opposed to using voice as any kind of input mechanism.

Many participants gave high marks to the system’s approach of mapping the horizontal location of their screen cursor to their physical location and the vertical location to their smartphone screen. However, technical difficulties in which some users had to repeatedly refresh the webpage on their phone due to lost WebSocket connections contributed to a certain level of annoyance.

One of the major usage challenges that many participants commented on was the appropriate appearance of screen prompts and other feedback at eye height. Designing user interfaces/feedback for large displays without blocking screen content is a continuing challenge for this type of research.

## DISCUSSION AND FUTURE WORK

In terms of overall performance, the results were very encouraging in regards to the usefulness of the overall interaction interface. Using standalone methods such as cross-device platforms or voice only methods have shown limited usability in the past [29]. However, in our case, we see that users adapt well, when multi-modal inputs; touch screen and voice are

used in conjunction with automatic interaction mechanisms based on spatial tracking.

Large display rich immersive spaces such as the one presented in this work draw significant amount of user attention. So, in designing interaction interfaces that do not overwhelm users, it is important to devise methods that require minimal attention. In this regard, using ubiquitous means such as smartphones, has shown considerable success in our work. Furthermore, allowing users to move freely and using the voice commands selectively, only for content generation, helped users to continuously focus on the screen and the task at hand instead of having to repeatedly glance at the phone screen or manually type input commands.

The multi-modal user interface presented in this work and its success has led us to work towards building use cases that go well beyond the game play experiments presented in this work. We are working towards building a language learning classroom use case, where students match language characters to images. Image selection and placement tasks based on combination of spatial intelligence, cross-device interaction and voice input in a large immersive space can be re-purposed to support classroom activities, where the room in itself is a teaching tool in contrast to conventional classrooms. The learning outcomes through student feedback and overall success of our interface will be important in furthering interaction design choices going forward.

While we only reported 2-user studies here, our immediate next step is to accommodate 3–6 users simultaneously, to fully realize the potential of our immersive environment as a multi-user space. In addition to direct extensions of the experiments we discussed here, we are investigating how the screen space for each user can be dynamically defined based on their location rather than constrained to one side of the screen.

We are also working to replace the worn Bluetooth microphones with an ambient microphone array that uses beamforming, along with the users’ known locations, to extract utterances for verbal input. Finally, we hope to conduct more systemic eye-tracking experiments to explore where the users look on the big screen and how often/under what circumstances they glance down at their phone “touchpad”.

## REFERENCES

1. Christopher Ackad, Andrew Clayphan, Martin Tomitsch, and Judy Kay. 2015. An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1227–1238.
2. Christopher Ackad, Martin Tomitsch, and Judy Kay. 2016. Skeletons and silhouettes: Comparing user representations at a gesture-based large display. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2343–2347.
3. Craig Anslow, Pedro Campos, and Joaquim Jorge. 2016. *Collaboration Meets Interactive Spaces*. Springer.

4. Matthias Baldauf, Florence Adegeye, Florian Alt, and Johannes Harms. 2016. Your browser is the controller: advanced web-based smartphone remote controls for public screens. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays*. ACM, 175–181.
5. Rafael Ballagas, Jan Borchers, Michael Rohs, and Jennifer G Sheridan. 2006. The smart phone: a ubiquitous input device. *IEEE Pervasive Computing* 1 (2006), 70–77.
6. 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*. ACM, 121–130.
7. Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. 2011. Code space: touch+ air gesture hybrid interactions for supporting developer meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 212–221.
8. Frederik Brudy, David Ledo, Saul Greenberg, and Andreas Butz. 2014. Is anyone looking? mitigating shoulder surfing on public displays through awareness and protection. In *Proceedings of The International Symposium on Pervasive Displays*. ACM, 1.
9. Xiang'Anthony Chen and others. 2014. Air+ touch: interweaving touch and in-air gestures. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM.
10. Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. Defanti, Robert V. Kenyon, and John C. Hart. 1992. The CAVE: Audio Visual Experience Automatic Virtual Environment. *Communication of the ACM* 35 (1992), 64–72. Issue 6.
11. Carolina Cruz-Neira, Jason Leigh, Michael Papka, Craig Barnes, Steven M. Cohen, Sumit Das, and others. 1993. Scientist in Wonderland: A Report on Visualization Applications in the CAVE Virtual Reality Environment. *IEEE Symposium on Research Frontiers in Virtual Reality*, 59–66.
12. Vishal Doshi, Sneha Tuteja, Krishna Bharadwaj, Davide Tantillo, Thomas Marrinan, James Patton, and G Elisabeta Marai. 2017. StickySchedule: an interactive multi-user application for conference scheduling on large-scale shared displays. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays*. ACM, 2.
13. Yuan Du, Haoyi Ren, Gang Pan, and Shjian Li. 2011. Tilt & touch: Mobile phone for 3D interaction. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 485–486.
14. Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, JD Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, and others. 2013. CAVE2: a hybrid reality environment for immersive simulation and information analysis. In *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 864903–864903.
15. Kirstie Hawkey, Melanie Kellar, Derek Reilly, Tara Whalen, and Kori M Inkpen. 2005. The proximity factor: impact of distance on co-located collaboration. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. ACM, 31–40.
16. Tom Horak, Ulrike Kister, and Raimund Dachzelt. 2016. Presenting business data: Challenges during board meetings in multi-display environments. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*. ACM, 319–324.
17. Ulrike Kister, Konstantin Klamka, Christian Tominski, and Raimund Dachzelt. 2017. GraSp: Combining Spatially-aware Mobile Devices and a Display Wall for Graph Visualization and Interaction. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 503–514.
18. Ulrike Kister, Patrick Reipschläger, and Raimund Dachzelt. 2016. Multilens: Fluent interaction with multi-functional multi-touch lenses for information visualization. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*. ACM, 139–148.
19. Regis Kopper, Mara G Silva, Ryan Patrick McMahan, and Doug A Bowman. 2008. Increasing the precision of distant pointing for large high-resolution displays. (2008).
20. Martin Kropp, Craig Anslow, Magdalena Mateescu, Roger Burkhard, Dario Vischi, and Carmen Zahn. 2017. Enhancing Agile Team Collaboration Through the Use of Large Digital Multi-touch Cardwalls. In *International Conference on Agile Software Development*. Springer, Cham, 119–134.
21. Christian Lander, Marco Speicher, Denise Paradowski, Norine Coenen, Sebastian Biewer, and Antonio Krüger. 2015. Collaborative newspaper: Exploring an adaptive scrolling algorithm in a multi-user reading scenario. In *Proceedings of the 4th International Symposium on Pervasive Displays*. ACM, 163–169.
22. Ricardo Langner, Ulrike Kister, and Raimund Dachzelt. 2019. Multiple Coordinated Views at Large Displays for Multiple Users: Empirical Findings on User Behavior, Movements, and Distances. *IEEE transactions on visualization and computer graphics* 25, 1 (2019), 608–618.
23. Can Liu. 2014. Leveraging physical human actions in large interaction spaces. In *Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology*. ACM, 9–12.
24. Petri Luojus, Jarkko Koskela, Kimmo Ollila, Saku-Matti Mäki, Raffi Kulpa-Bogossia, Tommi Heikkinen, and Timo Ojala. 2013. Wordster: collaborative versus competitive gaming using interactive public displays and mobile phones. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*. ACM, 109–114.

25. Shahzad Malik, Abhishek Ranjan, and Ravin Balakrishnan. 2005. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*. ACM, 43–52.
26. Andrea Nutsi. 2015. Usability Guidelines for Co-Located Multi-User Interaction on Wall Displays. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM, 433–438.
27. Andrea Nutsi and Michael Koch. 2015. Multi-User Usability Guidelines for Interactive Wall Display Applications. In *Proceedings of the 4th International Symposium on Pervasive Displays*. ACM, 233–234.
28. Markus Rittenbruch, Andrew Sorensen, Jared Donovan, Debra Polson, Michael Docherty, and Jeff Jones. 2013. The cube: a very large-scale interactive engagement space. In *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*. ACM, 1–10.
29. Amir E Sarabadani Tafreshi, Andrea Soro, and Gerhard Tröster. 2018. Automatic, Gestural, Voice, Positional, or Cross-Device Interaction? Comparing Interaction Methods to Indicate Topics of Interest to Public Displays. *Frontiers in ICT* 5 (2018), 20.
30. Gyanendra Sharma, Jonas Braasch, and Richard J. Radke. 2016. Interactions in a Human-Scale Immersive Environment: the CRAIVE-Lab. In *Cross-Surface Workshop at ISS2016*. Niagara Falls, Canada.
31. Alireza Sahami Shirazi, Christian Winkler, and Albrecht Schmidt. 2009. Flashlight interaction: a study on mobile phone interaction techniques with large displays. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 93.
32. Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM, 137–146.
33. Ulrich von Zadow and Raimund Dachzelt. 2017. Giant: Visualizing group interaction at large wall displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2639–2647.
34. James R Wallace, Ariel Weingarten, and Edward Lank. 2017. Subtle and Personal Workspace Requirements for Visual Search Tasks on Public Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6760–6764.
35. Katrin Wolf, Yomna Abdelrahman, Thomas Kubitz, and Albrecht Schmidt. 2016. Proxemic zones of exhibits and their manipulation using floor projection. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays*. ACM, 33–37.
36. Soojeong Yoo, Callum Parker, Judy Kay, and Martin Tomitsch. 2015. To dwell or not to dwell: an evaluation of mid-air gestures for large information displays. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*. ACM, 187–191.