Affordances of Input Modalities for Visual Data Exploration in Immersive Environments

Sriram Karthik Badam*
University of Maryland
College Park, MD, USA

Arjun Srinivasan†
Georgia Institute of Technology
Atlanta, GA, USA

Niklas Elmqvist‡
University of Maryland
College Park, MD, USA

John Stasko§
Georgia Institute of Technology
Atlanta, GA, USA

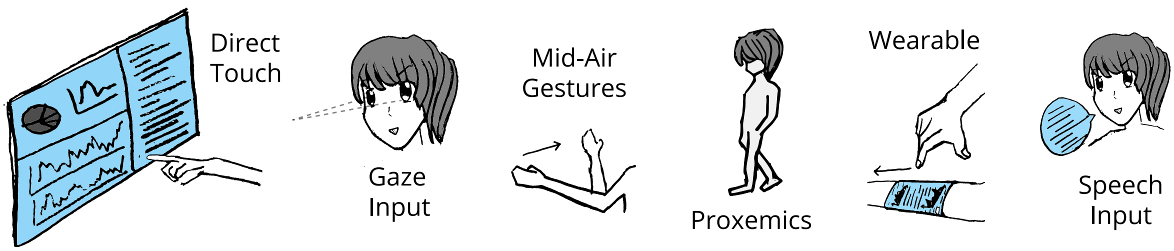

Figure 1: Input modalities for visual data exploration covered in this paper.

Abstract

There has been a consistent push towards exploring novel input, display, and feedback technologies for sensemaking of data. However, most visual analytic systems in the wild that go beyond a traditional desktop just utilize commercial large displays with direct touch, since they require the least effort to adapt from the desktop/mouse setting. There is a plethora of device technologies that are yet to be tapped. Through this paper, we want to bring attention to available modalities for input for visual exploration within immersive display environments. These can be environments that contain multiple wall and floor displays, or projections to create immersion. We first isolate the low-level interaction tasks performed by a user based on a usage scenario for visual exploration. We focus on egocentric visual exploration in these environments, and introduce input modalities that enable interactions directly between a human body and objects of interest on the interface without a mediator in the middle (e.g., a handheld smartphone). Based on this, we identify affordances of different input modalities—touch, speech, proxemics, gestures, gaze, and wearable—in terms of the interaction tasks from the envisioned scenario. Finally, we discuss how modalities can be combined to complement each other and leverage their advantages in the immersive environment. By doing so, this paper provides guidelines for new system developers to utilize the best input technologies for their immersive analytics applications.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Interaction styles;

1 Introduction

Developing an immersive analytics system is not straightforward. Beyond the monetary challenges, these systems need to be assembled to some extent, if not entirely, by coupling multiple technologies together. For instance, CAVE systems[1] can provide floor and wall displays within a room to create immersion, but to interact with them in the 3D space, motion capture platforms[2] are often used. However, the selected input technology may not provide the required freedom of expression in terms of interaction for visual exploration [25]. Case in point, gestural and proxemic interactions for visual analytic tasks have so far been only observed to be effective for simple interactions (e.g., changing zoom levels and switching between specific level of details in visualizations) [2, 11, 13]. Each input modality has specific affordances in terms of the interactions for visual analytic operations, and therefore, these affordances should be closely considered when developing new immersive visual analytics systems.

In this paper, we identify affordances of input technologies for supporting visual exploration. We focus on input modalities that enable *egocentric exploration* through interactions that are directly between the human body and objects of interest in the immersive environment, and not through an intermediate device (e.g., a handheld smartphone). To do so, we first discuss a potential usage scenario involving visual exploration in immersive room that contains wall displays. We extract specific low-level interaction tasks based on the scenario. Following this, we present the input modalities of interest—direct touch, speech, proxemics, gestures, gaze, and wearable input (seen in Figure 1) and discuss their affordances in terms of supporting the interaction tasks.

The affordances are discussed based on the freedom of expression of each input modality. For instance, touch input is very expressive and enables users to easily focus on specific visualizations in an interface, navigate through direct manipulation, and easily pick items of interest in a visualization for further exploration. However, creating visualizations using touch input requires interface options to capture the user's specification.[3] This freedom of expression based categorization led to Figure 2, which highlights the affordances for interaction tasks on a four-point scale for each modality. Based on this, it is apparent that speech and touch inputs can be coupled to best support all the interaction tasks. However, direct touch is only feasible when the user is close to a display. Other interesting and potentially useful couplings include, speech + mid-air gestures, proxemics + wearable input, and speech + gaze input, within the immersive environment. While the affordances are explained through a specific scenario, the principles can be applied generally to other immersive environments created by AR/VR headsets or even futuristic room-scale holographic displays.

*e-mail: sbadam@umd.edu
†e-mail: arjun010@gatech.edu
‡e-mail: elm@umd.edu
§e-mail: stasko@cc.gatech.edu

[1]CAVE: `http://www.visbox.com/products/cave/`

[2]VICON: `https://www.vicon.com/products/camera-systems`

[3]Previous interactive design tools [12, 20] helped convey the visualization specification in traditional desktop/mouse settings.

2 Envisioned Environment and Usage Scenario

The choice of target device(s) and the immersive environment created by them play a pivotal role in supporting immersive analytics. Considering a single environment with all possible input modalities and displays (touch-screen displays, projections, AR, VR, etc.) together is beyond the scope of one paper. To limit our scope to practical use cases, we envision the target setting for discussions in this paper to be a common thinking space such as board rooms with interactive surfaces or displays distributed in the room.

We imagine a visual interface in this immersive environment that presents a dataset through multiple visualizations. For the purpose of generality, we assume a simple visual design that exposes the data items within the dataset directly on the visual interface by utilizing the important visual variables (i.e., location, size, shape, and color). The individual visualizations use granular designs where each point in the visualization has a data context; for instance, a line chart with each point on the line representing an attribute from the dataset. This design is ecological since appropriate data transformations including aggregation and sampling can be used to create these point-based designs. Since each point in the visualization has a data context, the regions in space also have a context within the original dataset (e.g., corresponding to data attributes).

Usage Scenario. As an example of how such a setting may be used in practice, consider Sarah, an executive officer for a book publishing house. Sarah needs to make a presentation on the annual performance for the previous fiscal year to inform decisions about product lines and types of products (books) the company should focus on for the upcoming year. Sarah has the data for the previous fiscal year with details about individual sales, product type, region of sale, etc. To brainstorm about her upcoming presentation and explore this data freely, Sarah loads the dataset in an open space board room in her office with an interactive wall display and a comfortable couch. Below, we describe a usage scenario highlighting Sarah's experience in this setting. We discuss the scenario in terms of tasks to give a general overview. We discuss the possible input modalities and their affordances, and give examples of how various modalities could be used to support these tasks in the next section.

Settling herself comfortably on the couch, Sarah begins her exploration. She first creates a bar chart for the overall sales over quarters (**create**). To see the distribution of sales across product types, Sarah modifies the bar chart to a stacked bar chart showing sales by book genre (**reconfigure data mapping**). Noticing "fantasy literature" as the highest selling book type across quarters, she decides to explore it further. She highlights with color and annotates the bars for fantasy literature sales with the percentage of sales they contribute to in each quarter (**reconfigure graphical property**). Sarah adds a new map visualization to the wall showing distribution of sales for fantasy literature books around USA (**create, focus**). She notices that southern California and Florida have most sales. Intrigued by these hotspots, she looks more closely and notices that the hotspots are around Los Angeles, California and Orlando, Florida (**pick, find, filter**). Knowing that both locations are close to Walt Disney parks and Universal studios, she assumes that the higher sales attribute to an effect of park visits, as children get intrigured by park experiences and buy books to read new stories and engage with their favorite fictional characters.

She walks around thinking some more about what she could explore next and returns to sit on the couch, casually glancing through the visualizations she generated. Looking at the stacked bar chart (**focus, navigate**), Sarah notices that art books are not among the highest selling book types and wonders why since it seems logical to have them as an extension to fantasy literature. To identify the reasons, Sarah creates a scatterplot matrix showing the investment and profit for books across all genres (**create**). Next, Sarah compares the fantasy and arts genres in terms of investment-profit ratio (**reconfigure data mapping**). She notices that art books have

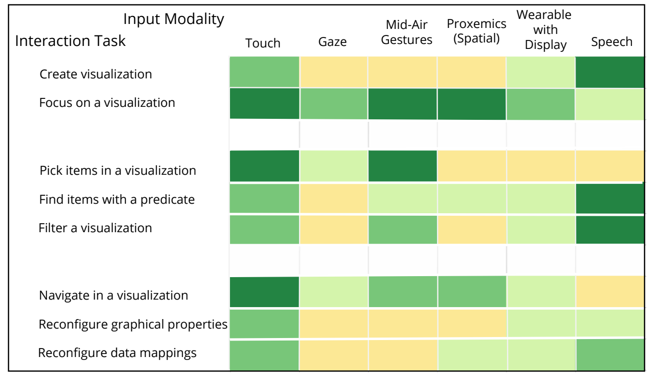

Figure 2: The affordances of various input modalities for each task on a four-point scale: not suitable directly ( ), requires additional interface options to be usable ( ), suited for simple cases of a task ( ), best suited for the interaction task ( ). The additional interface elements can be, for instance, a virtual cursor in gaze input to feedback of the focus, or UI widgets on a wearable smartwatch for filtering or changing graphical properties of a large display visualization.

a higher profit margin but have not been invested in by their company. On the other hand, fantasy literature books have a moderate investment-profit ratio and a large investment. She starts preparing her presentation based on these findings.

3 Tasks and Affordances

In the usage scenario, there are certain interaction tasks that are used by Sarah to visually explore the data in her enhanced office room. These tasks resemble the interactions specified in popular task taxonomies [25]. To identify the affordances of input modalities for supporting these tasks, we first define them more generally to extend them to a general taxonomy [25]. We then describe the user actions with each input modality to perform the task, along with the freedom of expression supported by the input modality for the particular task. We also discuss the aspects of fatigue, distance of action, and role of input modality in general for justifying the affordances. As mentioned earlier, we are interested in input modalities that create to egocentric exploration opportunities in the immersive environments, where the interactions happen between a human body and objects of interest without a mediator in the middle (e.g., a handheld smartphone) [19].

3.1 Interaction Tasks

The usage scenario covers eight interaction tasks.

- **Create** a new visualization on the interface.
- **Focus** on a specific visualization within the interface.
- **Pick** individual data items within a visualization.
- **Find** data items satisfying a predicate logic (e.g., $a < 5$).
- **Filter** selected data items by removing others.
- **Navigate** a visual with pan and zoom around focii.
- **Reconfigure graphical properties** in a visualization.
- **Reconfigure data mappings** driving a visualization.

These low-level interactions are connected to the high-level tasks defined by Yi et al. [25], and extend them to our scenario. For instance, pick and find are two ways to *select* content in a visualization. However, in contrast to Yi et al. [25], our list differentiates them since the cognitive effort for performing these two types of selections can be different from a user's perspective—forming predicates can be more complex than picking individual data items. Similarly, our reconfiguration tasks resemble the *encode*, *reconfigure*, and *abstract/elaborate* tasks from Yi et al. [25], but differentiate based on the graphical properties vs. data mappings. We admit that our interaction task taxonomy is by no means exhaustive or tailored towards immersive analytics, but we believe it offers a starting point to identify the affordances.

3.2 Affordances of Input Modalities

We are interested in six input modalities for visual exploration in our scenario. The capabilities of the input modalities in terms of the task affordances are broadly categorized in Figure 2. Here we introduce each modality and its input technology followed by its affordances for the tasks in terms of the freedom of expression.

3.2.1 Touch

Direct touch input with the displays in the environment can help directly interact with the visualizations. This input can be enabled, for instance, through capacitive sensors embedded within a display. Naturally, this input is only possible when very close to the target display. Touch interaction has three main actions: tap specific points in space, drag/move fingers in space, and gesture with single or multiple fingers. Due to these degrees of expression, touch is one of the most expressive modalities. Touch input can effectively support (1) focusing on specific visualizations in the environment by choosing them directly, (2) pick individual items in a visualization by tapping them, (3) navigating within a visualization by drag movements to zoom and pan, and also perform multi-focus navigation through multi-touch.

Other interactions can be performed with touch input, but need further interface support. To create new visualizations, specific interface features to specify the data attributes of interest and visual mappings by touch need to be present. To find data items based on predicates, additional interface widgets such as dynamic sliders and option menus are required. Touch gestures—e.g., a remove gesture—needs to defined to filter out uninteresting items in visualizations. Finally, reconfiguration interactions involving graphical items on a visualization—changing mark shapes or arrangement within a visualization—can be done directly through touch actions, but additional interface widgets are required to change color schemes, data mappings (e.g., replacing attributes), and data transformations (e.g., aggregation parameters).

3.2.2 Speech

Prior work has shown that users of visualization systems may be able to express their questions and intents more freely using natural language [1, 7], allowing them to perform a range of tasks [23]. In the envisioned scenario, speech input can allow the user to issue queries and interact with a visualization without being too close to the display or in its field of view. Speech queries can have a high freedom of expression, only bounded by the user's ability to express a query in natural language. Therefore they can effectively help the user, (1) create new visualizations by expressing which attributes and/or representations, or potentially, even the analytic intent, (2) find data items satisfying some predicate logic, and (3) filter the visualizations by expressing the intent to remove uninteresting items. With sufficient interface features for disambiguation [6, 21], speech input can also be used to reconfigure data mappings by specifying attributes to be added or removed (e.g., color based on book genre) and transformations on the underlying data attributes (e.g., show average values instead of counts). By designing appropriate language specifications and interface options, we can also support reconfiguration of graphical aspects (e.g., arrange attribute X in ascending order or change colors to a blue-red color scale), as well as focusing on specific visualizations in the interface.

Having said this, major design challenges in speech input arise for picking individual items from a visualization. This can be inconvenient unless there are sufficient and short labels that can be used to form a natural language query to pick multiple items. It is also hard to navigate to a specific region in a visualization, since spatial regions cannot be easily specified with natural language unless they have identifiable semantic aspects with respect to the dataset; otherwise, navigation can be arbitrary or even imprecise (e.g., zoom in on the top left).

3.2.3 Proxemics

Proxemic interaction refers to utilizing the spatial relationships [9] between people and artifacts in an environment including position, movement, distance, orientation, and identity. These variables can be obtained by using motion capture cameras within the environment. Interfaces built on these interactions automatically adapt to these user/device attributes [14]; for instance, a proxemic video player will start playing a movie when the user sits on the couch [3]. Due to the variety of proxemic dimensions, this input modality, representing spatial interaction, can be quite useful for visual exploration tasks. Within our usage scenario, proxemic input is best suited for interface-level operations such as focusing on a new visualization in the interface [2]. With some assumptions, proxemics can also be used for navigating within a visualization (pan and zoom) based on the attention of the user (e.g., orientation of their body) [13]. Furthermore, previous research has also considered manipulating the level of detail in visualizations—reconfigure data mappings—by using interaction zones in front of a display. For instance, when far from a display, a particular amount of visual content is shown [2, 11]. However, the remaining interactions in our scenario, are quite challenging to perform with just proxemic input. Finding data items that satisfy a predicate involves design considerations in terms of how to translate a user's field of view into a logical operation. Remaining interactions—create visualizations, pick individual data items, filtering, and reconfigure graphical properties—are hard to map naturally to proxemic dimensions making them unsuitable for this modality.

3.2.4 Mid-Air Gestures

Gestural input, especially mid-air, can help explicitly navigate the content shown in visualizations (e.g., zoom and pan) [2]. A long list of gestures designed for an interface may increase the freedom of expression of the user, making it possible to perform many interaction tasks. However, this is not a feasible choice as it complicates the user training and potentially confuses the users during the exploration. Hence we discuss the affordances based on two main types of gesture designs: (1) pointing gestures to convey interest in a specific unit in the interface, and (2) gestures utilizing hand movements based on the expected changes within a visualization [15]. Given these choices, gestures—similar to touch input—can effectively support (1) specifying the focus of the user within the entire interface and (2) picking specific items in the visualizations through pointing actions. With sufficient disambiguation and some assumptions, it is possible to filter visualizations and navigate within a visualization through gestural actions. Furthermore, with some interface options, simple predicates can be constructed for find interactions based on hand movements along axes of data attributes. However, this input is not suitable for reconfiguring graphical properties or data mappings as there will be complex design considerations in developing gestures that freely support them.

3.2.5 Gaze

Gaze input possible through eye tracking technologies offers potential opportunities to adapt visual interfaces based on the visual attention of the user. However, this also adds a challenge: using the visual channel for input as well as output can lead to unwanted changes to the content of the interface. For this reason, gaze input is typically based on the visual focus of the user's eyes as well as the duration of focus. Compared to other modalities, gaze supports fewer interactions. It is ideal for tracking the focus of the user within the entire interface. It can also be useful to pick specific items in a visualization and navigate (pan and zoom) towards the items in user's focus; however, the duration of gaze needs to be effectively configured to make the interaction seamless. Remaining interactions—create visualizations, find items based on predicates, filter, and reconfigure—are not suitable with gaze since the complexity of these interactions is too high to be easily supported by the freedom of expression of gaze input.

3.2.6 Wearable input device with a display

Wearables with a display attached to them such as smartwatches and on-body input devices [10] enhance an immersive environment with secondary interactive displays that are attached to the user's body to offer more freedom of expression in interactions. At the same time, wearable input modality falls into egocentric interaction by leveraging proprioception and eyes-free interaction for body-centric operations [22] (which is not possible with handheld devices). This modality also enables remote interaction from a distance from the immersive displays. With sufficient interface options on the wearable device, the focus of the user in terms of the visualization in the interface can be easily conveyed. For instance, a smartwatch can be used to perform swipe input to pick a visualization on a large display, without even looking at the watch. To support other interactions, the on-body/wearable interface needs to be designed effectively. For instance, to create new visualizations there should be options on a smartwatch interface to specify the attributes to visualize. Find, filter, navigate, and reconfigure interactions similarly require interface elements on the watch to convey the specific intent. Finally, picking individual data items through a wearable input device is not suitable since a responsive version of the large display interface should be placed on the small wearable display to pick the items; this has considerable limitations due to the mismatch in display size.

4 Discussion

Each input modality has specific expectations in terms of the system behaviors. In some cases, this is for reducing the amount of effort by being proactive; for instance, to automatically react to user's gaze rather than prompting them to gaze. Coupling the input modalities can also further enhance the visual exploration by overcoming individual tradeoffs. Here we discuss these aspects.

4.1 Combining Modalities: Why? When? How?

As stated earlier, Figure 2 discusses affordances of various input modalities individually. However, prior work in the broader HCI community [17, 18], and even recent work within the visualization community [2, 24] has shown that it is natural and effective to use multimodal input. Coupling input modalities can create more design opportunities, where advantages of one modality can be used to overcome the challenges of another (e.g., the ambiguity in speech can be corrected by precise selections offered via touch).

Coupling modalities is not straightforward, however. Users may prefer different modalities for different tasks and the usage patterns of modalities may vary as well (e.g., sequential vs. simultaneous multimodal interaction). One of our primary goals with Figure 2 was to present an overview of strengths and weaknesses of various modalities with respect to specific tasks in a scenario such as the one presented earlier. A direction for future work is to consider how two or more of the discussed modalities, that are suited based on Figure 2, can be combined to facilitate a full spectrum of tasks and foster a more fluid interaction and visual analysis experience for the users [5].

Exploring preferences for input modalities based on tasks, understanding how people combine modalities, and elucidating if preferences change for different combinations of modalities (e.g., is speech always the most commonly used input modality? do people use it less when combined with wearable technology as compared to touch?) are all open questions. For example, contrary to common assumption, prior work has shown that when using speech and gesture, people rarely perform "put-that-there" [4] style interaction using both modalities simultaneously [16]. We hope discussions and ideas presented in this article help future systems consider which modalities can be mixed to foster an enhanced visual analysis workflow, and encourage them to identify the most effective input combinations for their specific visual exploration tasks.

4.2 System Behavior: Proactive vs. Reactive

Most existing work in the visualization community exploring post-WIMP input has considered reactive system behavior and how systems can respond to "intentional" user interactions. However, in a setting such as the one presented in the usage scenario earlier, some input modalities are potentially more suited for proactive system behavior by default whereas others may be more suited for reactive behavior. For instance, proactively adjusting the size of the visualization in focus based on the user's proximity to the display can help preserve the focus and allow the user to look at the same set of data from different distances (e.g. while sitting or moving around in the scenario discussed earlier). Reactively doing so demands additional effort from the user, which is not ideal for longitudinal analytics sessions. On the other hand, proactively listening to user utterances and adapting to them would lead to an interruptive and frustrating user experience.

With some input modalities, proactive behavior can be more suited for certain tasks and less for others. For instance, proactively updating the "focus" visualization or resizing a visualization based on a user's location (proxemics) may be more effective than navigating the visualization content (zoom/pan within a visualization). This was identified in previous work to combine proxemic (implicit/proactive) and gestural interactions (explicit/reactive) [2] into a hybrid interaction model for visual exploration. Additionally, in a setting with multiple modalities, depending on the sequence of use of the modalities, proactive behavior in response to one modality may interfere with the input/output for another. Preserving the user's workflow and identifying when to activate/deactivate proactive behavior is also an open challenge when mixing modalities for visual exploration. For instance, proxemics may be well suited for identifying shifts in the focus (or target) visualization when there are multiple visualizations to choose from. When combining speech and proxemics, systems would need to identify when to apply proactive behavior and shift focus *versus.* when not to do so based on the user's movement and potential changes in global coherence in the dialog between the user and the system [8].

Acknowledgments

This work was supported by the U.S. National Science Foundation award #1539534. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

References

[1] J. Aurisano, A. Kumar, A. Gonzales, K. Reda, J. Leigh, B. Di Eugenio, and A. Johnson. "Show me data": Observational study of a conversational interface in visual data exploration. In *IEEE VIS '15 (Poster paper)*, 2015.

[2] S. K. Badam, F. Amini, N. Elmqvist, and P. Irani. Supporting visual exploration for multiple users in large display environments. In *Proc. VAST'16*. IEEE, 2016.

[3] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pp. 121–130, 2010.

[4] R. A. Bolt. "Put-that-there": Voice and gesture at the graphics interface. In *Proceedings of the 7th Annual Conference on Computer*

Graphics and Interactive Techniques, SIGGRAPH '80, pp. 262–270, 1980.

[5] N. Elmqvist, A. V. Moere, H.-C. Jetter, D. Cernea, H. Reiterer, and T. Jankun-Kelly. Fluid interaction for information visualization. *Information Visualization*, 10(4):327–340, 2011.

[6] T. Gao, M. Dontcheva, E. Adar, Z. Liu, and K. G. Karahalios. Datatone: Managing ambiguity in natural language interfaces for data visualization. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 489–500. ACM, 2015.

[7] L. Grammel, M. Tory, and M.-A. Storey. How information visualization novices construct visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943–952, 2010.

[8] B. J. Grosz and C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational linguistics*, 12(3):175–204, 1986.

[9] E. T. Hall. *The Hidden Dimension*. Anchor Books, Garden City, NY, 1966.

[10] C. Harrison, D. Tan, and D. Morris. Skinput: appropriating the body as an input surface. In *Proceedings of the ACM SIGCHI conference on human factors in computing systems*, pp. 453–462, 2010.

[11] M. R. Jakobsen, Y. Sahlemariam Haile, S. Knudsen, and K. Hornbæk. Information visualization and proxemics: design opportunities and empirical findings. *IEEE Trans. Vis. Comput. Graphics*, 19(12):2386–2395, 2013.

[12] N. W. Kim, E. Schweickart, Z. Liu, M. Dontcheva, W. Li, J. Popovic, and H. Pfister. Data-driven guides: Supporting expressive design for information graphics. *IEEE transactions on visualization and computer graphics*, 23(1):491–500, 2017.

[13] U. Kister, P. Reipschläger, F. Matulic, and R. Dachselt. Bodylenses: Embodied magic lenses and personal territories for wall displays. In *Proc. ITS '15*, pp. 117–126. ACM, 2015.

[14] N. Marquardt and S. Greenberg. Informing the design of proxemic interactions. *IEEE Pervasive Computing*, 11(2):14–23, 2012.

[15] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay. Mid-air pan-and-zoom on wall-sized displays. In *Proc. of the ACM Conference on Human Factors in Computing Systems*, pp. 177–186, 2011.

[16] S. Oviatt. Ten myths of multimodal interaction. *Communications of the ACM*, 42(11):74–81, 1999.

[17] S. Oviatt. Multimodal interfaces. *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*, 14:286–304, 2003.

[18] S. Oviatt and P. Cohen. Perceptual user interfaces: multimodal interfaces that process what comes naturally. *Communications of the ACM*, 43(3):45–53, 2000.

[19] T. Pederson. Egocentric interaction. In *Workshop on What is the Next Generation of Human-Computer Interaction*, pp. 22–23, 2006.

[20] A. Satyanarayan and J. Heer. Lyra: An interactive visualization design environment. In *Computer Graphics Forum*, vol. 33, pp. 351–360. Wiley Online Library, 2014.

[21] V. Setlur, S. E. Battersby, M. Tory, R. Gossweiler, and A. X. Chang. Eviza: A natural language interface for visual analysis. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 365–377. ACM, 2016.

[22] G. Shoemaker, T. Tsukitani, Y. Kitamura, and K. S. Booth. Body-centric interaction techniques for very large wall displays. In *Proceedings of the ACM Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pp. 463–472, 2010.

[23] A. Srinivasan and J. Stasko. Natural language interfaces for data analysis with visualization: Considering what has and could be asked. In *Proceedings of EuroVis '17*, pp. 55–59, 2017.

[24] A. Srinivasan and J. Stasko. Orko: Facilitating multimodal interaction for visual network exploration and analysis. In *IEEE Transactions on Visualization and Computer Graphics*, 2018.

[25] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Trans. Vis. Comput. Graphics*, 13(6):1224–1231, 2007.