

VisConductor: Affect-Varying Gestural Widgets for Animating Dynamic Data Visualization in Augmented Video Presentations

by

Temiloluwa Femi-Gege

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2023

©Temiloluwa Femi-Gege 2023

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis represents a culmination of research in study design, application development, data analysis, and academic writing, enriched by the guidance and advice of two pivotal figures in my academic journey.

As the primary researcher and writer, my responsibilities spanned the conceptualization and execution of the research. This entailed developing the specific software application under study, designing the research framework, recruiting and interviewing participants, performing detailed data analysis, and synthesizing the findings into the final thesis document.

Dr. Jian Zhao, as my supervisor, provided comprehensive advice and feedback that were instrumental across all facets of the research process. His insights were vital in refining my methodologies and ensuring the academic rigor of the work.

Matthew Brehmer, in his capacity as an expert advisor, offered critical contributions that enriched each stage of the project. His guidance was particularly valuable in deepening our understanding of the research area.

The combined support and insights from Dr. Jian Zhao and Matthew Brehmer were essential to the successful completion of this thesis.

Abstract

To enhance presentations involving data in remote settings, we introduce VisConductor, an authoring and presentation tool that allows presenters to manipulate data visualizations overlaid on their webcam feed with hand gestures. Contrasting traditional platforms like PowerPoint, which distance presenters from their content, VisConductor fosters a more interactive and engaging experience. VisConductor equips presenters with the ability to configure gestures for control over expressive visualization animations, foreshadow visualization transitions, animate the disclosure of annotations, and accentuate data points during remote presentations. These gestures are integrated into configurable widgets, allowing presenters to command content by executing gestures within widget boundaries. Gesture widgets also offer real-time hand pose feedback during live presentations. To evaluate VisConductor, we conducted two user studies focusing on presenters ($N = 11$) and audience members ($N = 11$). Feedback highlighted the capabilities of VisConductor to facilitate interactivity and engagement in remote data visualization presentations, offering valuable insights into the future development of gesture-based presentation tools.

Acknowledgements

First and foremost, I wish to express my profound gratitude to Dr. Jian Zhao, my esteemed supervisor at the University of Waterloo. His unwavering support and mentorship throughout my graduate journey have been invaluable. Dr. Zhao's timely and insightful suggestions, coupled with his expert guidance on courses and research topics, have been instrumental in the realization of this thesis. What I particularly cherish is the balance he struck, granting me the freedom to delve into areas of my own interest while steering me with his seasoned perspective.

Equally deserving of special recognition is Matthew Brehmer, whose extensive discussions, invaluable insights, and deep experience in my research domain have been a beacon of light. Alongside Dr. Zhao, Matthew's commitment was exemplified in our weekly meetings, where our brainstorming sessions played a pivotal role in molding this thesis to its present form.

I would be remiss if I did not extend my heartfelt appreciation to my lab colleagues. Their collective wisdom, constructive critiques, and suggestions have significantly enriched this work. Beyond the academic contributions, their active participation in various studies added vibrancy to my graduate school experience, making it as fulfilling as it was educational.

Furthermore, I'd like to convey my sincerest appreciation to the University of Waterloo. Its esteemed faculty, state-of-the-art facilities, and the conducive academic environment have been indispensable in shaping my academic pursuits. The lessons learned and experiences gained here will forever remain etched in my memory.

In closing, this thesis, is truly a testament to the collective efforts and encouragement of these remarkable individuals and institutions. Their influence, both direct and indirect, has indelibly marked every page, and for that, I am eternally grateful.

Table of Contents

Author's Declaration	ii
Statement of Contributions	iii
Abstract	iv
Acknowledgements	v
List of Figures	viii
List of Tables	xi
1 Introduction	1
2 Related Work	4
2.1 Fostering Engagement with Data Visualizations	4
2.2 Communicative Use of Gestures	5
2.3 Gesture Controlled Presentation Interfaces	6
3 Designing VisConductor	9
3.1 Design Considerations	9
3.2 Usage Scenario	11

4	VisConductor System	14
4.1	Hand Gestures	14
4.2	Affect Varying Gestural Control	16
4.3	Authoring Widgets	18
4.4	Presentation Delivery Experience	19
4.5	Implementation Details	19
5	Evaluation	21
5.1	Presenter Study	21
5.2	Audience Study	22
5.3	Observations and Participants' Reflections	23
5.3.1	Utility	23
5.3.2	Effectiveness	24
5.3.3	Customizability and Usability	25
5.3.4	Challenges	26
5.4	Audience Engagement	27
6	Discussion	29
7	Limitations and Future Work	32
8	Conclusion	34
	References	35

List of Figures

1.1	A presenter is using VisConductor to perform a remote live presentation with their audience regarding world life expectancy and income. The left shows the presenter view whereas the thumbnail on the right shows what the audience sees. While the audience sees the primary content, the presenter is aided with additional widgets to ensure precise gesture control. (A) Gesture Guide Widget: A delineated rectangle outlines the specific boundary within which the presenter’s hand gesture should be executed. (B) Gesture Recognition Feedback: Upon detecting the presenter’s hand pose, a progressively expanding green circle emerges, functioning as a visual timer. This informs the presenter of the duration for which the pose should be maintained for accurate detection.	3
2.1	User Interface of VisConductor: (A) Presentation Window: Central display area showcasing the presenter’s webcam feed with integrated widgets such as gesture guides, annotations, and charts. (B) Timeline Slider: Allows for previewing specific moments on the visualization timeline. (C) Stories Tab: Lists the distinct segments of the presentation. (D) Widget Settings Tab: Area for customizing the currently selected widget. (E) Widget List: Organized list of available widgets layered within the presenters feed. (F) Chart Widgets: Tools for visual data representation. (G) Annotation Widgets: Instruments for explanatory notes or highlights. (H) Gesture Widgets: Modules for gesture recognition and feedback.	8

3.1	Presenter view within VisConductor featuring various interactive elements: (A) Open Hand : Reveals the title through a text annotation within the Presentation Window. (B) Pistol : Highlights specific countries within the data visualization. (C) Rectangle Framing : Activates the foreshadowing of future states of chart elements. (D) Dialling : Initiates playback by tracking the stroke-based motion.	11
3.2	Overview of widgets in VisConductor. (A1) General Gesture Widget Settings : Allows presenters to set duration for poses, specify dominant hand, label widgets, and define gesture operation mode. This mode choice determines the subsequent forms shown in A2-A5. (A2) Selection Settings : Appears when mode is set for highlighting. Configures chart entities to be highlighted. (A3) Foreshadowing Settings : Shown when gestures are set to foreshadow animations. Enables selection of entities to foreshadow and the mode of foreshadowing. (A4) Playback Settings : Configures animation aspects like duration, SVG morphs, and easing for playback and assigning them to an affect. (A5) Annotation Settings : Allows presenters to set the revealed annotation upon gesture execution. (B) Annotation Widget Settings : Fine-tunes annotations in terms of color, opacity, reveal duration, and reveal easing. (C1) Chart Widget Settings : Provides a field for JSON data input and offers configuration for visualization. (C2) Chart Widget Preferences : Adjusts keyframes, item opacities, and foreshadowing graphics for animations.	13
4.1	Supported hand gestures in VisConductor: (A) Pistol, (B) Open Hand, (C) Rectangle Framing, (D) Range Framing, and (E) Dialling.	15
4.2	Foreshadowing techniques for the two chart types in VisConductor. (A) Scatter Plots - Point Foreshadowing : Depicts initial and final position of the entity. (B) Scatter Plots - Trajectory Foreshadowing : Illustrates the entity's path towards its future state. (C) Bar Chart Race - Point Foreshadowing : Highlights the future position of the entity/bar. (D) Bar Chart Race - Trajectory Foreshadowing : Utilizes a bump chart to trace the bar's positions, showcasing its movement, e.g., Peru's ascent from 6th to 5th place over three days.	17

5.1	Distribution of responses on the engagement questionnaire of the audience study, where 1 = “strongly disagree” and 5 = “strongly agree”. The red dashed line indicates the median and the shaded region denotes the interquartile range (IQR).	27
-----	--	----

List of Tables

4.1	Overview of gesture types and characteristics in VisConductor.	14
-----	--	----

Chapter 1

Introduction

The surge in remote work and education necessitates the development of presentation tools that enhance remote communication, allowing speakers to convey information in an engaging and interactive manner [8, 74]. In this thesis, we build on recent work considering the power of hand gestures that accompany speech to enhance remote presentations, since hand gestures remain a fundamental aspect of non-verbal communication [71, 76]. They enrich our communicative repertoire, allowing for a more holistic and nuanced exchange of ideas and emotions. Whether used intentionally or spontaneously, they add depth and dimension to our conversations.

During remote presentations, the audience often faces the dilemma of dividing their attention between the presenter and the visual aids. This division can hinder the cohesive flow of communication and understanding, as attendees oscillate between focusing on the speaker’s discourse and the provided slides. Integrating visual aids directly into a presenter’s video feed, can seamlessly combine the presenter and the content, allowing presenters to better focus the attention of their audience using their gestures thereby enhancing the presentation experience [29]. In particular, presentations about data, benefit from the use of visualizations as they are an influential and engaging medium for information delivery [54, 22]. During these presentations, presenters are frequently confined by the primary interaction mechanism—using various peripherals to click and trigger predefined actions such as animated transitions or motion effects. Such click-based interactions, though functional, can sometimes limit the dynamism and interactivity that visualizations can potentially offer. This limitation is even more pronounced in remote settings as existing tools operate on click-based interactions, devoid of features that leverage natural hand gestures for both interaction and communication. Causing a disconnect between the speaker and their content.

In addition to considering the benefits of overlaying visual aids in enhancing remote presentations, it is essential to recognize the power of animation in visual storytelling. When showing visualizations, using animations is one key strategy to heighten a presentation by eliciting affective responses and retaining audience attention [37]. Situated in animated visualizations, presenters manipulate the playback, manage the animated display of different kinds of annotations, select/highlight elements, and even borrow cinematic techniques such as foreshadowing, which can greatly elevate the effectiveness and engagement of their presentations. Yet, presenters planning to explain data using animations are confined to using peripherals to control visualizations. This confinement to peripherals removes the invaluable communicative utility of gestures, especially in remote settings, creating a palpable disconnect. While software like mmhmm [68], Zoom [20], WebEx [66], and Apple’s presenter overlay [31] allow for the visual compositing of presenter video and visual aid, presenters cannot manipulate visual aids without programming expertise.

In light of these challenges, we introduce *VisConductor*, an augmented presentation authoring tool that allows presenters to control visualization animations via hand gestures (Figure 1.1). *VisConductor* aims to streamline the authoring process of presentations overlaying visual aids on the video feed, while amplifying their interactivity and engagement with gesture control. Presenters can advance the temporal state of a visualization, trigger the animated display of annotations, and foreshadow the future state of visualizations. These can be achieved with gesture widgets, which not only provide visual cues during presentation delivery, but also simplify the authoring process, this being a limitation of prior research. The gestures incorporated into *VisConductor* are deliberately chosen, encompassing a balance of deictic and iconic gestures, which are foundational in presentations [59]. To evaluate *VisConductor*, we conducted two user studies focusing on presenters ($N = 11$) and audience members ($N = 11$), respectively. In the presenter study, participants were asked to prepare a presentation with *VisConductor* using a pre-defined script, story, and dataset, whereas in the audience study, participants were asked to experience a live remote presentation given by the researchers. The results suggest that *VisConductor* provides an effective authoring and presenting experience for presenters as well as enhances audience engagement and retains their interest. To summarize, our contributions encompass:

- An authoring tool, VisConductor, for creating and remotely presenting visualizations with animations, annotations, and foreshadowing, all controllable through hand gestures, to boost audience engagement.
- A set of gesture widgets designed in VisConductor, which are tailored to authoring the gesture control of visualizations, also double as live presentation guides to enhance overall usability.
- Comprehensive user studies assessing the effectiveness and usefulness of VisConductor, providing crucial insights for the future development of gesture-based presentation tools.

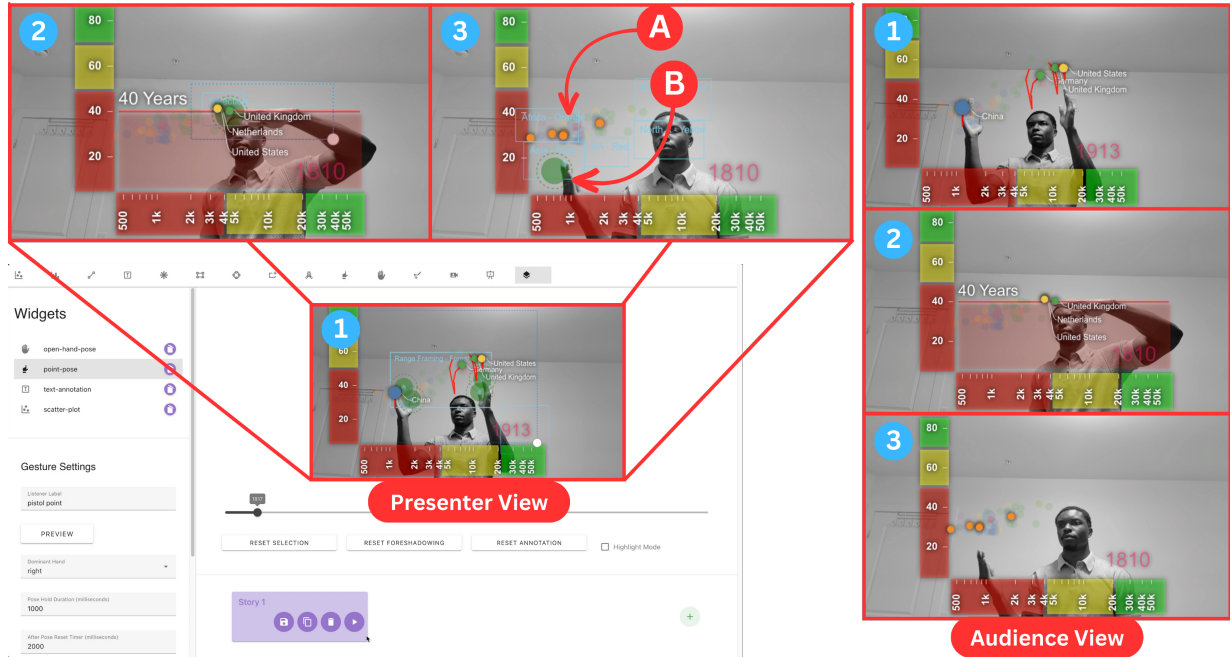


Figure 1.1: A presenter is using VisConductor to perform a remote live presentation with their audience regarding world life expectancy and income. The left shows the presenter view whereas the thumbnail on the right shows what the audience sees. While the audience sees the primary content, the presenter is aided with additional widgets to ensure precise gesture control. **(A) Gesture Guide Widget:** A delineated rectangle outlines the specific boundary within which the presenter’s hand gesture should be executed. **(B) Gesture Recognition Feedback:** Upon detecting the presenter’s hand pose, a progressively expanding green circle emerges, functioning as a visual timer. This informs the presenter of the duration for which the pose should be maintained for accurate detection.

Chapter 2

Related Work

2.1 Fostering Engagement with Data Visualizations

Data visualizations are instrumental in translating complex datasets into impactful narratives for presentation and communication [30, 34, 58]. The strength of data visualizations as a communication tool extends beyond mere representation. First, this includes the strategic use of animation [55], which can enhance audience engagement and comprehension. Second, annotations act as a vital guide, enhancing the audience’s understanding by directing their attention to the most crucial aspects of the data [63]. Whether applied to axes, data points, or trend lines, annotations can be data-driven or manually added, serving as narrative markers in the visualization [53].

This affective dimension of storytelling is further emphasized through animation. While animation can effectively demonstrate changes over time or evolving patterns, it also introduces an emotional narrative to the visualization [55, 3, 13, 37]. This emotional connection is particularly vital in presentations, as it can enhance audience engagement and information retention [23]. Amini et al. [3] highlights that pictographic representations in charts can stimulate viewer engagement by evoking emotions, given their appealing nature compared to standard charts. The introduction of animations to these pictographs further amplifies this effect. Lan et al. [37] also discussed the influence of animation speed on the affective impact of chart animations, suggesting that it is an important factor to consider in the design process. Another innovative concept in data visualization is visual foreshadowing, which serves as a potent strategy for audience engagement [39]. By providing a glimpse of the future state of an animation, it prepares the audience for what lies ahead, enhancing comprehension and engagement.

There exists a plethora of libraries and tools that enable the development of animated visualizations, with D3 [12] and Plotly [52] serving as general-purpose visualization libraries, gganimate [51] as a specialized library focusing on animation, and Data Animator [70] as a dedicated tool for authoring visualization animations. However, they come with their own set of limitations. It takes a considerable amount of effort to implement a solution and none of them support gestural animation control out of the box. As a result, the finished product from using these tools is predominantly confined to traditional click-based interactions. This limitation hinders the potential for a more immersive and interactive data visualization experience. Tools like DataClips [5] and Roslingifier [59] have emerged to facilitate the creation of data videos. Though adept at creating data-driven stories, they remain click-based and are restricted to scatter plot visualizations. In contrast, VisConductor enables the authoring of expressive, data-driven animated scatter plots and bar chart race animations. Moreover, it empowers presenters to perform operations on visualizations using gestures in real-time during presentations, providing a more interactive experience.

2.2 Communicative Use of Gestures

According to Goldin-Meadow [26], gestures are a multifaceted tool in communication, often conveying subtleties and nuances that spoken words might miss. They act as visual complements, sharpening the message’s clarity and making it more accessible and memorable for the audience. This visual support not only heightens audience engagement but also fosters better retention of the information shared. Supporting this notion, Tieu et al. [71] delved into the amplifying effect of gestures on spoken language especially when presenting complex or abstract ideas. Zeng et al. [76] further emphasized the importance of apt gestures in diverse communication settings, from public speaking to daily interactions. Their findings suggest that for communication to be truly effective, the gestures used must resonate with the spoken content. Kang et al. [32] also touched upon the potency of gestures, especially when they are in perfect sync with the ideas they aim to represent. This synchronization often emerges naturally; speakers, during their presentations, tend to use gestures either as a conscious strategy or as a spontaneous reaction to their narrative.

There are various types of gestures, e.g., deictic, iconic, metaphoric, and beat gestures [33, 18]. Deictic gestures (e.g., pointing) help direct attention; iconic gestures visually represent specific ideas or objects; and metaphoric gestures symbolize abstract concepts. Carstens et al. [18] conducted a study to identify common presentation gestures, noting that one-handed gestures are often deictic, while two-handed gestures tend to be iconic or metaphoric. Our tool aims to incorporate these recognized gestures for both functional

and communicative purposes.

2.3 Gesture Controlled Presentation Interfaces

In conventional presentation settings, studies indicated that bare hand gestures were favored compared to a laser pointer [16]. While presenters frequently employ gestures, these gestures often exist in a space separate from the animations or visual aids, leading to a perceived spatial disjointedness. The divide becomes even more pronounced in online presentations where a distinct presenter camera feed is utilized, making the integration of gestures with visual aids a challenging endeavor. A quintessential example of combining data-driven visualizations with gestures is Hans Rosling’s “200 countries, 200 years” presentation [11]. Rosling employed an animated scatter plot and various types of gestures to craft a compelling narrative. However, it worth noting that his presentation was not as spontaneous as it appeared. His actions and gestures were meticulously choreographed, and visual aids were integrated during post-production, rather than being controlled in real-time, creating an illusion of seamless interaction with the visualizations.

Existing tools have attempted to create platforms where users can use gestures for communication and functional control of visual aids in real time [29, 41, 40]. They have demonstrated the ability of gesture control to significantly enhance the interactive experience for presenters while simultaneously fostering greater engagement and connection with the audience. These solutions typically follow the pattern of a webcam or camera feed overlaid with interactive visuals. Developments in real-time pose estimation technologies such as Kinect [60], OpenPose [17], and MediaPipe [72] have opened up opportunities to explore the potential of gesture-controlled visuals in remote settings. Saquib et al. [56] introduced an application allowing for the mapping of gestures to control dynamic graphics, which provides users with a wide range of creative freedom that requires some amount of preparation and supports real-time interaction with graphics at performance-time. However, a notable limitation of their work is the confined support for data-driven animated visualizations. While users can map gestures to alter the states of the visualization, these state alterations lack the sophistication of animated transitions. This limitation implies a discernible gap in the field where presenters are impeded in their ability to craft truly animated visualization transitions and to employ animations to elicit affect effectively. Hall et al. [29] explored the use of gestures to interact with visualizations dynamically; they focused less on authoring these interactions but instead looked into a finite predefined set of gestures that are discussed as deictic and also functional. However, they omitted the gestural control over multi-state animations; the support for animation is primarily

centered on the revelation of annotations.

During live performances, a significant challenge is the mental overload faced by the presenter [56]. The first source of this overload stems from the need to remember which gestures to perform, when to execute them, and anticipating the graphical effects they will produce. Charade’s research emphasized the need for clear feedback in gesture-based presentation systems to aid presenters [8] as they highlight that gestures are not self-revealing. These issues are not limited to presentations but also extend to broader AR/VR environments [48, 7], where designing both the environment and interaction space can be complex, especially for non-experts. Therefore, there is still a need to explore more solutions, especially when working with data-driven visualizations. Presenters, especially when dealing with extensive timelines like Hans Rosling’s presentation [11], must manually track the current timeline state, ensuring they do not overshoot or undershoot their target. While Saquib et al.’s system [56] provides configurable transitions for other animations and offers the convenience of dragging and dropping data into the interface to create visualizations, its handling of data-driven chart animations lack transitions and results in clunky changes between states. VisConductor aims to address these challenges with modular gesture widgets, allowing users to seamlessly integrate, configure, and manage various components of the presentation with gesture control, during both authoring and presenting.

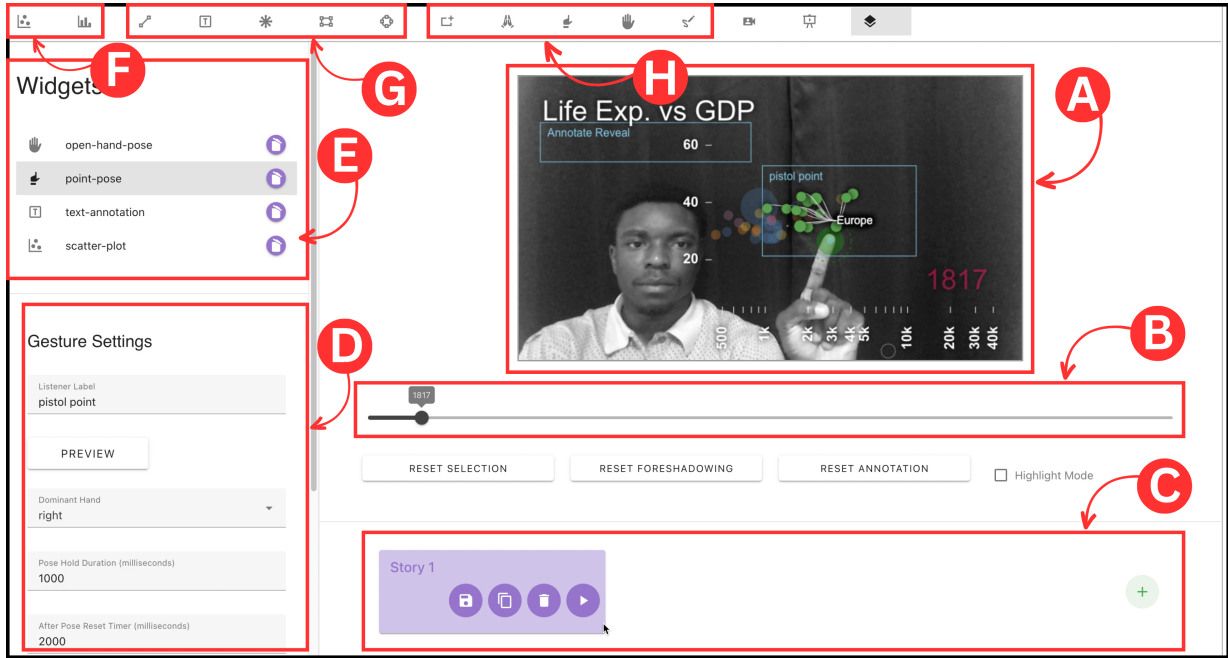


Figure 2.1: User Interface of VisConductor: (A) **Presentation Window**: Central display area showcasing the presenter’s webcam feed with integrated widgets such as gesture guides, annotations, and charts. (B) **Timeline Slider**: Allows for previewing specific moments on the visualization timeline. (C) **Stories Tab**: Lists the distinct segments of the presentation. (D) **Widget Settings Tab**: Area for customizing the currently selected widget. (E) **Widget List**: Organized list of available widgets layered within the presenters feed. (F) **Chart Widgets**: Tools for visual data representation. (G) **Annotation Widgets**: Instruments for explanatory notes or highlights. (H) **Gesture Widgets**: Modules for gesture recognition and feedback.

Chapter 3

Designing VisConductor

This section introduces the design of VisConductor, followed by a short usage scenario to demonstrate the key features of the tool.

3.1 Design Considerations

Our objective was to optimize online presentation with an emphasis on enhancing data visualization experiences. We crafted a series of mock-ups by employing a “wizard of oz” approach synchronizing gestures with predefined animated sequences to simulate gesture-controlled animations. We then conducted a series of interviews and co-design sessions with a communication expert who has over two decades of experience in strategic communications, leadership development, and global philanthropy (referred to as E1). E1 has personally trained over 10,000 individuals across various sectors and regions, which provided invaluable insights into our project. Drawing upon E1’s expertise, we integrated essential feedback, leading to a set of design considerations.

D1: Provide a communicative yet simple set of hand gestures. Gestures should not require undue effort or a heavy mental load as this risks distracting the presenter from their communicative intents [21]. The system should be responsive, recognizing gestures in real time [21]. Gestures should also be designed to naturally convey meaning to enhance the presentation experience. As Hall et al. point out, there is a need to balance the richness of expressive body language with operational gestures, those which are specifically designed to control or manipulate the visual aids [29]. Gestures need to be intuitive, and the selection process should prioritize gestures that have a clear relationship with the functionalities

they execute, further reducing the cognitive load on the presenter [21]. Confirmed by E1, the system should adopt a set of gestures that presenters commonly use, and these gestures should be intentionally chosen to be communicative while ensuring they enhance, rather than diminish, the presenter’s natural expressiveness.

D2: Support diverse animation-based operations on visualizations. The literature survey underscores the multifaceted potential of data visualizations in enhancing comprehension and engagement during presentations [63, 53, 3, 55, 39], which informed this design goal. Specifically, the manipulation of animations should be supported, which not only represent data changes but evoke emotions, solidifying their importance in presentations. As illuminated in the work done by Amini et al. [3], affective animations infuse an emotional narrative, intensifying the bond with viewers. Moreover, echoed by E1, presenters need foreshadowing, a concept endorsed by Li et al. [39], which offers viewers a preview of what’s to come, amplifying both comprehension and engagement. Finally, annotations, whether they are data-driven or manually incorporated, have been identified as pivotal narrative markers in visualizations [63, 53]. Thus, the tool should allow for animated revealing of animations, which serves as a critical guide, augmenting the audience’s understanding by spotlighting data facets.

D3: Facilitate authoring gestural control over visualization operations. Overlaying visualization directly with the presenter’s webcam feed not only bridges the spatial gap between gestures and visual aids but also paves the way for a more immersive and integrated presentation experience. Thus, an effective authoring environment is required to permit presenters to craft the interactions and define the graphical effects resulting from these interactions [56, 29]. This goal is multifaceted: to offer a familiar interface that leverages existing presentation paradigms as well as to allow presenters a foresight into their authored interactions, ensuring they can envision, rehearse, and understand the flow of their presentation without facing undue mental strain. As E1 aptly noted, even the most intuitive gestures cannot compensate for poorly designed visualizations. By facilitating this clarity in authoring, we aim to allow presenters to focus on their content and delivery, rather than being preoccupied with the mechanics of their visual aids.

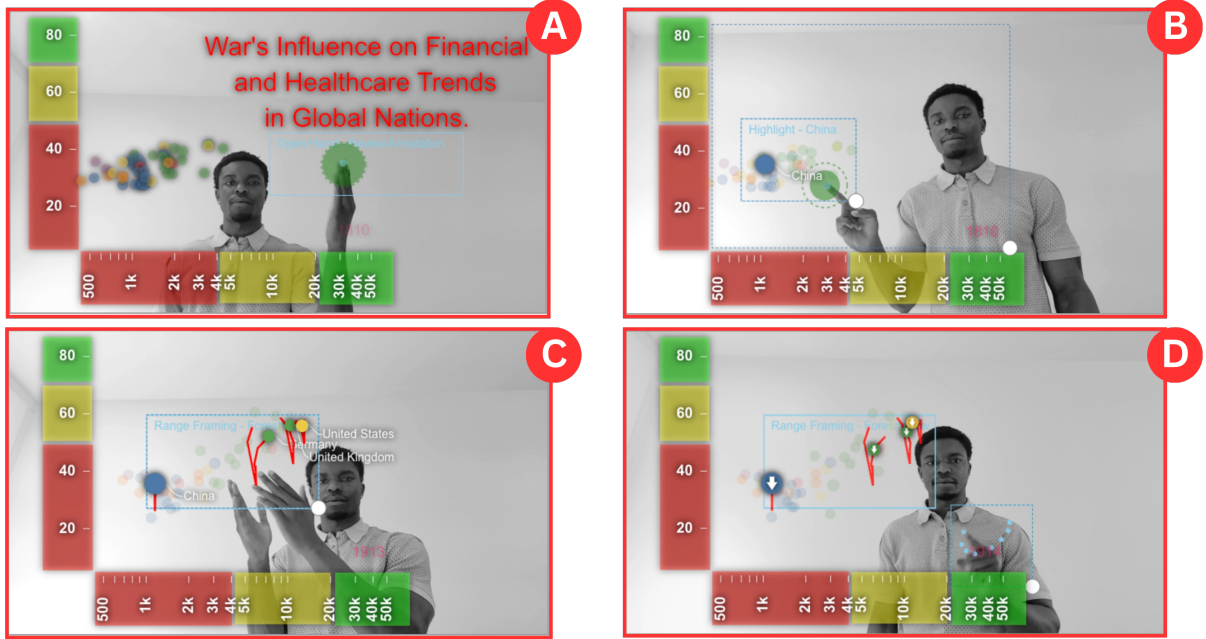


Figure 3.1: Presenter view within VisConductor featuring various interactive elements: (A) **Open Hand**: Reveals the title through a text annotation within the Presentation Window. (B) **Pistol**: Highlights specific countries within the data visualization. (C) **Rectangle Framing**: Activates the foreshadowing of future states of chart elements. (D) **Dialling**: Initiates playback by tracking the stroke-based motion.

3.2 Usage Scenario

Suppose that John is a public health worker, and he is preparing an essential remote presentation about the shifts in monetary and healthcare metrics across countries of the world during different time periods, visualized as an animated scatter plot. After launching the tool, John is presented with a graphical interface (Figure 2.1) to author his presentations with visualizations and gesture control over them.

In his first slide, John's primary intent is to set the scene. He integrates a scatter plot visualization into his first slide, feeding in his JSON dataset and fine-tuning the visualization details as shown in Figure 3.2-C1. To introduce the title of his presentation to the audience, he adds a text annotation labeled "War's Influence on Financial and Healthcare Trends in Global Nations." To underscore the gravity of his topic and its implications,

John chooses a slow, deliberate reveal for his text annotation. This decision mirrors the profound consequences of war, ensuring the message resonates with the audience.

Intending to emphasize aspects of the visualization, John integrates a communicative set of hand gestures into his slide. By configuring the *Open Hand* pointing gesture, John intends to direct the audience’s attention to specific areas, ensuring that as he points, corresponding animations dynamically unveil, thereby adding depth to his narrative and guiding viewers through the data.

In his subsequent slide, John’s objective is to draw the audience’s focus towards specific nations of interest. By leveraging the *Pistol* pointing gesture, he guides viewer attention to these main countries. When he points using this gesture, it not only acts as a directional cue but also amplifies the visual prominence of the highlighted nations. The resulting visual impact is showcased in Figure 3.1-B.

Anticipation is the theme of John’s next slide. Employing the *Range Framing* gesture around a cluster of countries he intends to foreshadow, John ensures a dual outcome: guiding the viewer’s attention and kick-starting the foreshadowing effect to showcase the future state of these countries before the animation plays. This mechanism mirrors the one depicted in Figure 3.1-C.

Lastly, John aims to play through the visualization states and elicit an affective response from his audience. He configures the *Dialing* gesture, to be performed around the displayed date on the visualization. This setup gives the illusion that John is manipulating time as he performs the dialing gesture. He then configures the animation to morph countries into downward arrows during the animation, echoing war’s devastating impacts, in line with the playback settings demonstrated in Figure 3.2-A4. The culmination of his work on this slide is presented in Figure 3.1-D.

As John prepares for his presentation, he shifts to the presenter view in VisConductor, granting him real-time feedback on his gestures as outlined in Figure 1.1. This perspective, invisible to his audience, enables John to seamlessly integrate his gestures—executing the *Open Hand* gesture to dynamically reveal the annotation, performing the *Pistol* pointing gesture to draw focus to specific nations, utilizing the *Range Framing* gesture to foreshadow clusters of countries, and employing the *Dialing* gesture to animate through the visualization states—ensuring his delivery is coherent, engaging, and immersive, with every gesture serving its intent and complementing the narrative.

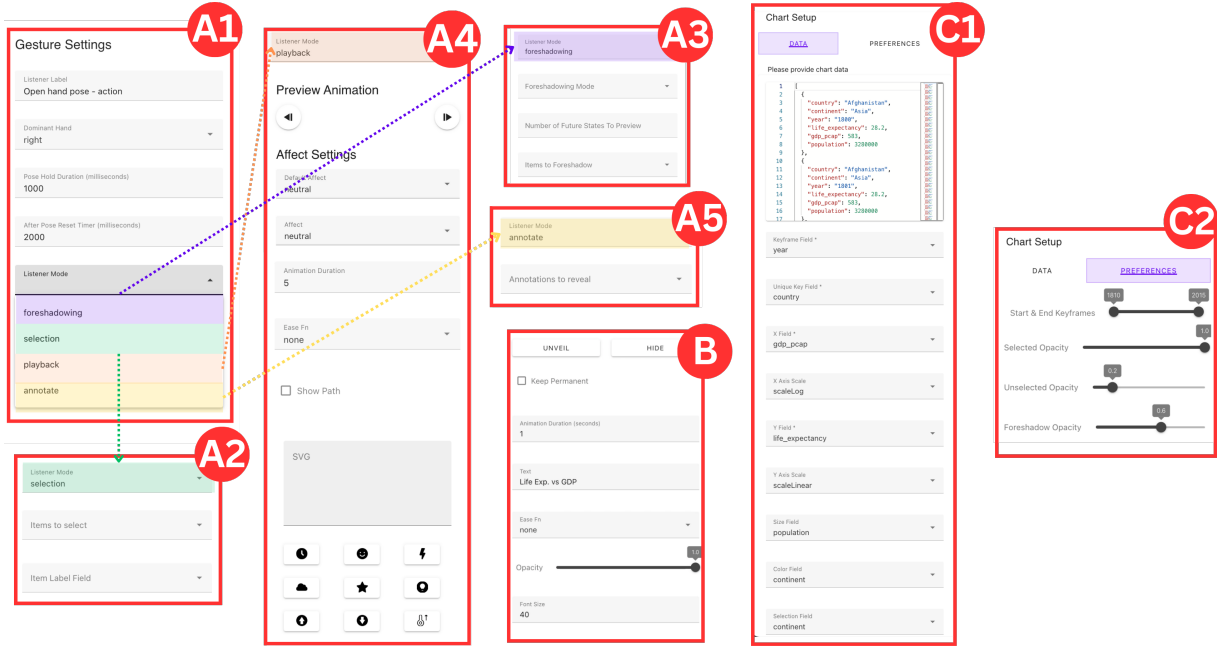


Figure 3.2: Overview of widgets in VisConductor. **(A1) General Gesture Widget Settings:** Allows presenters to set duration for poses, specify dominant hand, label widgets, and define gesture operation mode. This mode choice determines the subsequent forms shown in A2-A5. **(A2) Selection Settings:** Appears when mode is set for highlighting. Configures chart entities to be highlighted. **(A3) Foreshadowing Settings:** Shown when gestures are set to foreshadow animations. Enables selection of entities to foreshadow and the mode of foreshadowing. **(A4) Playback Settings:** Configures animation aspects like duration, SVG morphs, and easing for playback and assigning them to an affect. **(A5) Annotation Settings:** Allows presenters to set the revealed annotation upon gesture execution. **(B) Annotation Widget Settings:** Fine-tunes annotations in terms of color, opacity, reveal duration, and reveal easing. **(C1) Chart Widget Settings:** Provides a field for JSON data input and offers configuration for visualization. **(C2) Chart Widget Preferences:** Adjusts keyframes, item opacities, and foreshadowing graphics for animations.

Chapter 4

VisConductor System

Based on the aforementioned design goals, we developed VisConductor, each key component of which will be detailed below. The name mirrors an orchestra conductor’s hand gestures directing instruments, analogous to how VisConductor manages gesture and visualization widgets.

4.1 Hand Gestures

According to **D1**, our selection of gestures was constrained to the set of gestures that align with natural co-speech body language, often requiring little to no learning curve for users. Second, the chosen gestures should not place undue physical or cognitive burdens on the presenter. Third, the gestures must be communicative in nature, either deictic, iconic, metaphoric, or beat gestures to emphasize certain points. Last but not least, the number of distinct gestures should remain small.

Table 4.1: Overview of gesture types and characteristics in VisConductor.

Gesture	Characteristics		
Pistol	Pointing	Unimanual	Static
Open Hand	Pointing	Unimanual	Static
Range Framing	Shaping	Bimanual	Static
Rectangular Framing	Shaping	Bimanual	Static
Dialling	Pointing	Unimanual	Dynamic

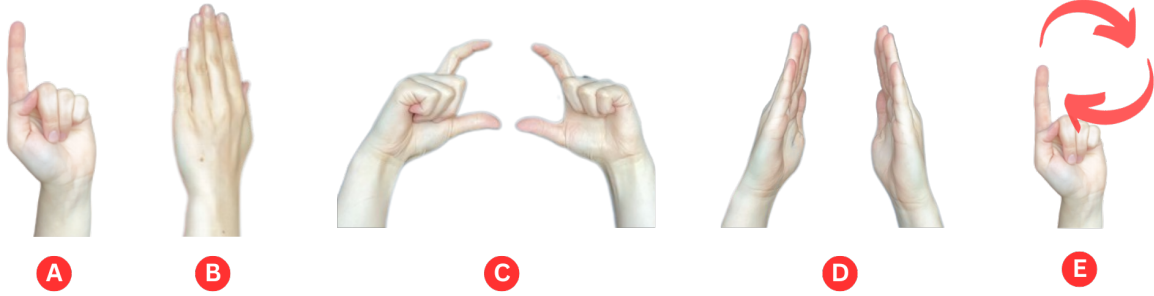


Figure 4.1: Supported hand gestures in VisConductor: (A) Pistol, (B) Open Hand, (C) Rectangle Framing, (D) Range Framing, and (E) Dialling.

Our gesture space design draws from the techniques and intentions identified by Shin et al. [59]. To enhance our understanding, we also examined the fact-pod videos produced by the Gapminder foundation [25], keenly observing the commonly employed hand poses in data video explanations. Overall, VisConductor supports five different gestures as illustrated in Figure 4.1 and Table 4.1. There exist three main design dimensions classifying these hand gestures. First, *pointing* is used to draw attention to specific data entities, while *shaping* conveys ranges and boundaries [59]. *Pistol* (Figure 4.1-A) and *Open Hand* (Figure 4.1-B) were selected for our pointing gestures, according to Carstens et al.’s classification [18] that accentuated the frequent use of these gestures. These gestures direct attention and emphasize specific components in VisConductor. Second, gestures are either *uni-manual* or *bi-manual*. We thus extended the *Pistol* with a slight bend in the index finger, creating a rectangular shape, which is symbolic in certain data contexts, referred to as *Rectangle Framing* (Figure 4.1-C). We also extended the *Open Hand* to its bi-manual version, *Range Framing* (Figure 4.1-D), which can denote a data range. These bi-manual gestures are also shaping gestures, primarily iconic or metaphoric in nature, embodying a visual or symbolic representation of the ideas they are intended to convey. Lastly, gestures can be *static*, where the pose is merely held, or *dynamic*, which requires motion tracking. Following **D1**, we only extended the *Pistol* to *Dialling* (Figure 4.1-E), which offers a metaphorical representation, symbolizing the progression of time for navigating timelines or traversing historical data. The dynamic versions of other gestures do not trigger any visualization operations, however, presenters can repetitively perform them at multiple locations, including *Open Hand*, *Rectangle Framing*, and *Range Framing*, to signify emphasis in VisConductor.

4.2 Affect Varying Gestural Control

VisConductor allows presenters to use hand gestures to control various operations on animated visualizations (D2).

VisConductor currently supports: the bar chart race and the dynamic scatter plot, and we have outlined plans to incorporate additional animated visualization types in future iterations, including animations for line charts and pie charts.

First, *Animation Playback* can be manipulated using gestures, where the state and progression of the chart is based on the underlying dataset provided. With respect to the affect-varying aspect of animations, VisConductor presents an advanced configuration framework (Figure 3.2-A4). This framework not only allows for meticulous visual storytelling but also promotes the communication of appropriate affects. Presenters have the freedom to shape animations by choosing from various ease functions and setting SVG shapes and icons as morph targets during playback. They can either utilize predefined SVGs or introduce their unique SVG data, offering notable flexibility. By *Affect Varying*, we refer to presenters’ ability to map specific ease functions and morph pairings to three emotional categories: positive, negative, and neutral. By adjusting these templates, presenters can preset the emotional tone of their animations. For example, in Figure 3.1-D, to play out an animation where chart elements move downwards, symbolizing the devastating loss of lives during World War I, and then make a return upwards, representing the economic recoveries that followed the war. The deliberate slow start but fast finish of the ‘ease-in cubic’ function serves to immerse the audience in the profound depth of the loss before offering a beacon of hope with the swift upward movement. To further emphasize this narrative, a morph transition to a downward arrow is employed, capturing the descent into the somber reality of wartime casualties.

Second, *Foreshadowing Reveal* is supported via a dedicated set of Foreshadowing Settings (Figure 3.2-A3), where presenters can choose between *point* and *trajectory* foreshadowing modes, both of which present future data states. Point foreshadowing juxtaposes a data point’s current state with its future one. This juxtaposition is symbolized using a distinct pulsing effect, as illustrated in Figure 4.2-A and Figure 4.2-C. On the other hand, the trajectory mode sketches out the path that a data point is set to embark upon. It is worth noting that while scatter plots lay out this journey as a fluid path, the bar chart race reveals an ephemeral bump chart as foreshadowing (Figure 4.2).

Third, *Annotation Reveal* supports both *data-centric* and *non-data-centric* annotations. For the former, through the Selection Settings (Figure 3.2-A2), presenters have the leverage to decide which chart components to spotlight through full opacity and which

ones to subtly fade out using reduced opacity. These decisions are not arbitrary; they are informed by identifiers within the dataset, ensuring that the presenter’s narrative remains data-driven. In contrast, independent from the chart data, presenters have the power to punctuate their narrative with textual or visual explanations. Configured through both the Annotation Settings (Figure 3.2-A5) and the Annotation Widget Settings (Figure 3.2-B), these annotations can be invoked through the hand gestures supported, amplifying the clarity and impact of the narrative.

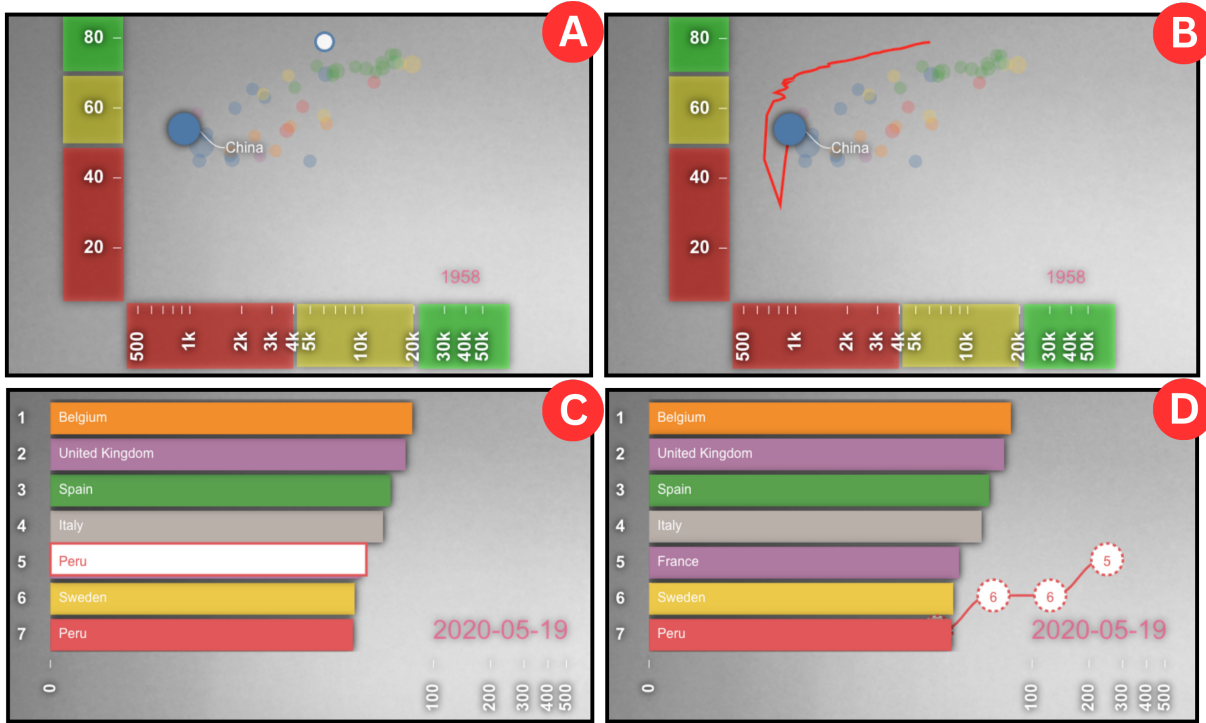


Figure 4.2: Foreshadowing techniques for the two chart types in VisConductor. **(A) Scatter Plots - Point Foreshadowing:** Depicts initial and final position of the entity. **(B) Scatter Plots - Trajectory Foreshadowing:** Illustrates the entity’s path towards its future state. **(C) Bar Chart Race - Point Foreshadowing:** Highlights the future position of the entity/bar. **(D) Bar Chart Race - Trajectory Foreshadowing:** Utilizes a bump chart to trace the bar’s positions, showcasing its movement, e.g., Peru’s ascent from 6th to 5th place over three days.

4.3 Authoring Widgets

In VisConductor, we embrace a familiar design environment reminiscent of presentation tools such as PowerPoint and Keynote (Figure 2.1), a VisConductor presentation consists of a set of *stories*, which correspond to the traditional concept of slides, offering a canvas for the presenter. This interface reduces the learning curve. Central to the design of VisConductor is the concept of *widgets* (Figure 3.2). Everything, from visual representations to interactive gestures, is encapsulated as a widget. This modular approach allows presenters to seamlessly integrate, configure, and manage various components within a story (D3).

The *Chart Widgets* display either of the two chart types supported in VisConductor: scatter plots and bar charts. Figure 3.2-C1 depicts the inputs the presenter has to provide after selecting a chart widget. They configure the dataset, animation states/keyframes, unique data point identifiers across keyframes, axis values, scales, and potential grouping of items. Other configurations include item opacity, start and end keyframe values to dictate where animations start and finish, and selection criteria (either by unique key or group identifier). Crucially, a story can contain only one visualization widget to maintain clarity.

While labels for chart entities are data-driven and applied on selection, the *Annotation Widgets* presented here are distinct. These widgets allow for the placement of non-data-driven textual and graphic components, akin to elements in traditional presentation tools. Each widget-based annotation can be dynamically revealed with customizable animations, encompassing specifications for duration, animation ease function, colors, and opacity (Figure 3.2-B).

The *Gesture Widgets* capture the essence of gestural interaction in VisConductor, as introduced in Figure 4.1. Each gesture widget offers tailored configurations such as pose detection duration, re-detection intervals, and the choice of which hand the pose should be detected for. Moreover, every gesture can be mapped to a specific visualization operation (Section 4.2). Once linked, additional settings related to the chosen feature emerge as shown in Figure 3.2-A1:A5. In a nutshell, the process is: add the desired widgets to a story, configure their individual settings, and then, for the gesture widgets, map them to desired chart animation, foreshadowing, or annotation reveal.

The system offers users the flexibility to assign any of the five gesture widgets to any of the four visualization operations, empowering gestures to activate these operations. However, we advocate some general guidelines to ensure that gestures resonate with the conveyed meaning of the operation, which are built as the defaults in VisConductor. For instance, the uni-manual pointing gestures, such as the *Open Hand* and *Pistol* gestures,

are inherently deictic in nature. Deictic gestures, serve to point out or draw attention. In VisConductor, we effectively harness the viewer’s instinct to follow the direction of the point, and employ these gestures to *Annotation Reveal*, ensuring the revealed items stand out, and *Foreshadowing Reveal*, indicating what is to come. For bi-manual gestures, their inherent deictic nature makes them well-suited for several other operations. The *Range Framing* and *Rectangle Framing* gestures, by virtue of their ability to encompass or demarcate a specific shape or area, are inherently designed to steer the audience’s attention to a particular section of the visualization. Paired with the *Annotation Reveal*, they ensure that a specific range or cluster of data points is accentuated prominently. Lastly, the *Dialling* gesture tracked using the stroke listener widget, stands out as the optimal choice for controlling *Animation Playback*. Its rotational motion inherently signifies the progression through time evoking the movement of an analog clock or wrist watch, making it a metaphorical gesture aligning with the concept of moving forward in a timeline, offering viewers a dynamic journey through the data’s temporal aspects.

4.4 Presentation Delivery Experience

Akin to the aforementioned tools, VisConductor encapsulates both a *presenter view* and an *audience view* (Figure 1.1). Compared to the audience view, the presenter view is enriched with gesture widgets—blue bounding boxes with labels, solely visible to the presenter, delineating the bounds for gesture execution. Upon performing a pose within these bounds, an expanding green bubble appears, filling a circle, and acting as a timer indicating the duration the pose needs to be held. This dual-view approach ensures a seamless presentation experience while retaining the necessary controls and feedback for the presenter.

4.5 Implementation Details

VisConductor operates within a web browser environment. By accessing the device’s webcam, it captures the live video feed and funnels this data directly to our gesture detection mechanism. Following successful gesture recognition, the system then actuates the relevant animations on the visualization canvas. This streamlined process, from webcam capture to visualization manipulation, ensures real-time responsiveness and a smooth user experience. The gestural detection is facilitated through the collaborative strengths of MediaPipe [72] and Finger Pose [57]. MediaPipe captures and forwards hand landmarks, which are pivotal for the accurate identification of gestures. Finger Pose is used for discerning specific finger

configurations, leveraging the landmarks supplied by MediaPipe. Together, they recognize and translate gestures into system commands, initiating a plethora of functionalities that enrich the tool’s overall interactivity. For the animation control, GreenSock (GSAP) [27], a high-performance JavaScript animation library, is used to enhance the dynamic visual elements on the canvas with fluidity.

Chapter 5

Evaluation

Following Hall et al.’s methodology [29], we conducted two separate studies of VisConductor: one centered on the presenter’s experience of authoring and presenting using the tool, and the other considered the audience’s perspective of engaging with a presentation crafted through VisConductor. Since there is no existing tool that offers similar functionality with respect to authoring live presentations of animated data visualizations controlled using hand gestures, we did not compare VisConductor against a baseline system. Moreover, Hall et al.[29] already verified that the audience has a more engaged experience with overlaying visualization on the video feed, compared to traditional screen sharing. Therefore, we decided to prioritize the qualitative aspects of the tool’s utility rather than a quantitative assessment of task completion relative to a baseline experience.

5.1 Presenter Study

The focus of this study was to understand the utility and usability of VisConductor with respect to presentation authoring and delivery.

Participants: We recruited 11 participants, 3 females and 8 males, aged 18–34, via mailing lists and word of mouth. They had a range of technical backgrounds including engineering, business, and computer science, of which 8 held post-graduate degrees and 3 were college graduates. Regarding technological proficiency with presentation applications, 5 self-identified as “Advanced,” 4 as “Expert,” and 2 as “Intermediate.” Regarding presentation skills, 4 self-identified as “Advanced,” and 7 as “Intermediate.” All participants had experiences with traditional in-person presentations, while 3 were familiar with giving

online presentations. As for the frequency of delivering presentations, 8 reported giving about one presentation a month and 3 reported doing so weekly. Their presentations incorporated various elements, including data visualizations, textual slides, images, diagrams, and multimedia elements such as videos.

Procedure: After signing the consent form and filling out the pre-study questionnaire regarding their backgrounds, we gave participants a walkthrough of VisConductor, highlighting its features and encouraging questions. Next, we introduced them to a template presentation with five stories, each embedded with a particular visualization widget: 5 participants encountered the bar chart race while 6 experienced the animated scatter plot, by assigning participants to one of the visualizations alternatively. Following the template, participants engaged in different activities, such as adding and adjusting charts, annotations, and gesture widgets, familiarizing themselves with VisConductor without constant guidance, though we provided assistance when needed. After, participants crafted a presentation using the tool, based on a provided script and stories to present. Lastly, we invited participants to a semi-structured interview. The whole study lasted about 1 hour and each participant received \$20 as remuneration.

5.2 Audience Study

The primary objective of this study was to assess the tool’s efficacy in presenting visualization with affect varying animations, visual foreshadowing, and highlighting, with a specific focus on how VisConductor might result in greater understanding and engagement.

Participants: We recruited another 11 participants, 4 females and 7 males, aged 18–44. Their technical backgrounds included engineering, computer science, and business. Of all the participants, 5 held post-graduate degrees and 4 were college graduates, the remaining 2 completed colleges. When assessing their proficiency in understanding data visualizations, 6 self-identified as “Advanced,” 4 as “Intermediate,” and 1 as “Beginner,” on a 5-point Likert Scale. Most (9 out of 11) frequently attended presentations, with 1 attending daily, 1 monthly, and 7 weekly. All had participated in in-person presentations, while 6 had attended online presentations.

Procedure: After signing the consent form, participants completed a pre-study questionnaire capturing their past experiences related to presentations. Then, participants watched a live presentation conducted in VisConductor over Zoom. While the content varied depending on the data visualization method—either bar chart race ($N = 5$) or animated scatter plot ($N = 6$)—we ensured to showcase all of VisConductor’s essential features, which are applicable across different visualizations. After watching it, participants

needed to provide a summary of the content, which helped us gauge their understanding and the tool’s ability to convey data narratives. Finally, participants were invited to share their experiences in a semi-structured interview and a questionnaire regarding their engagement during the presentation. The whole study lasted about 30 minutes for each participant and they received \$15.

5.3 Observations and Participants’ Reflections

Here we report findings from both studies, where “P#” denotes a presenter study participant and “A#” indicates an audience study participant.

5.3.1 Utility

Presenters appreciated having gestural control over animations. P2 mentioned, *“I think they help to develop the whole slide gradually, which makes the audience engage in the presentation.”* Adding depth to this perspective, P7 likens using VisConductor to a performance, stating, *“It gives the idea that I have more control over my slides. Almost like a performance for the audience.”* This reinforces the notion that VisConductor can serve as an instrument for crafting a riveting presentation narrative. Further, P2 underlined the interactivity of the tool for the presenter, mentioning that *“I think it can be intuitive for presenters... The presenter is actually working and doing physical gestures rather than just talking through the slides. It might be engaging for them as well.”* Segueing into the comparative advantages of VisConductor, P10 honed in on VisConductor’s strength in configuring gestural interactions, explaining the relative difficulty to achieve this in other platforms, *“The ease of interaction is what’s the best point... for the simple presentation tool like PowerPoint, it is harder to go within the presentation itself to point to stuff, play or annotate stuff... it takes much more technical expertise.”* This observation suggests that VisConductor addresses the challenges presenters may face with other platforms, bridging the gap between visualization operations and gestural interactions.

The general sentiment amongst audience members was that there is a need and desire for tools that allow for the creation and control over animated visualizations in presentations. A7 delved into the practical concerns regarding the setup of such tools, articulating: *“a concern would be how much effort does this take to set up? What level of abstraction do I deal with or is it video editing?”* This suggests part of the appeal of VisConductor lies in its ability to simplify the creation of animated data visualizations while being able to

control these animations using gestures. Building on these, another compelling dimension of VisConductor’s utility was brought up. A9 shed light on the significant aspect of visual attention control, implying how VisConductor enables presenters to effectively direct the audience’s focus: “... *this way, you can control what you want the audience to see... especially if you try to convey something emotional.*” This perspective adds a layer of depth to the discussion, emphasizing not just the tool’s technical capabilities but also its potential for affective resonance in presentations.

5.3.2 Effectiveness

To understand the effectiveness of message delivery using VisConductor, we asked about three gestural control features: affect-varying animation, visual foreshadowing, and annotation revealing.

Affective Animation: The use of animations was generally well-received. A10 highlighted the intuitive nature of certain animations that aided in understanding the data’s trajectory: “*there was like an arrow or something that showed up when I am expecting the the the circle to drop... I feel like these small things really help me understand the point you’re trying to convey.*” This specific animation—where elements morphed into downward-facing arrows as a prelude to their subsequent drop—served as an affective visual cue, emotionally priming the viewer for the impending event. Participants positively received the tool’s capability to integrate animations with gesture-based interactions. A11 noted, “*The animation... caught my attention because you were playing it...*” A9 added that the tool aids in showcasing affect-varying gestures alongside the graphics: “*helps if you want people to see your emotional gesture...makes it more clear to show which part of the presentation you want people to focus on whether it’s the graphics or is it both you and the graphics*” In essence, the integration of animations with gesture-based interactions not only enhances the comprehension of data but also accentuates the affective aspect of a presentation.

Visual Foreshadowing: Participants’ responses underpinned the importance of gesture-based interactions in conjunction with visual foreshadowing in enhancing engagement, feeling that the presenter’s presence behind the data emphasized important aspects of the visualization before playback. “*Your face being behind the data or being with the data definitely emphasizes the importance of where to look... especially the World War One part where you show the trend lines before the actual animation plays out. That was super interesting.*” -A7 From the presenter’s side, P4 felt that foreshadowing helped both in highlighting parts of the graph and piquing the audience’s curiosity, while P3 mentioned that it keeps the audience “*on their toes.*”

Annotation Revealing: Building on the insights on visual foreshadowing, another salient feature that emerged was the role of highlighting and selection of chart elements (i.e., annotations) in animated visualizations. When queried about its influence on comprehending data, responses indicated a positive impact. A7 recognized the effectiveness of color schemes and gestures in differentiating data and explaining its components: *“... your gestures explaining the axes and the different colors and how they move... It was easy to understand. I don’t think there were any gaps or lack of understanding because of too much information on the plot itself.”* From the presenter’s perspective, the benefits were also evident. P1 elaborated on the ease of use, mentioning, *“The gestures are good, I can just highlight which position of the graph I want to tell to the audience and we can set beforehand which part I will show so it will highlight accordingly.”* P4 further attested to the utility of these features, noting the synergistic effect of selection and foreshadowing in emphasizing specific graph segments: *“I mean the selection part, even the foreshadowing I really like those two. Helps me highlight certain parts of the graph. Just to show whatever is coming next.”* Presenters also discussed the competitive advantages of the annotation revealing over conventional presentation tools; they were vocal about the enhanced interactivity offered. P10 succinctly captured this sentiment, stating, *“The ease of interaction is what’s the best point for me at least because for the simple presentation tool like PowerPoint, it is harder to go within the presentation itself to point to annotate stuff.”* This observation suggests that the nature of this presentation method, coupled with its interactive capabilities, offers a more engaging and intuitive experience compared to traditional presentation tools.

5.3.3 Customizability and Usability

VisConductor’s ability to upload data, configure animations, adjust gesture widgets, and feedback mechanisms were central to this discussion. P3’s feedback was overwhelmingly positive regarding the authoring experience, emphasizing VisConductor’s customizability: *“authoring is pretty easy, it’s really accommodative...there are a lot of options to choose from.”* P9 appreciated the visual effects and customization capabilities emphasizing the ability to morph chart elements into various SVGs: *all the visual effects are really good...we can customize the icons...* However, P11 mentioned that while the tool offers more design options for animations, it requires greater attention to design details: *you got more design for the animation it means that you need to put more details into the design part...but once you get familiar with the whole system it allows you to create the perfect slides.*

On the topic of gesture widgets and feedback from these widgets, drawing inspiration from existing tools, P3 felt that VisConductor was familiar due to the widget-based ap-

proach. The user interface of VisConductor seemed intuitive to several users, reminiscent of other presentation authoring tools they have interacted with. This familiarity was a boon for users like P3, who found the transition to VisConductor smoother: *“it was easy for me as a first-time user to do it really fast...It felt familiar because when we use various other apps, we see these options available as widgets.”* Additionally, VisConductor’s flexibility and adaptability were evident in its customizability features. P1 appreciated the ability to manually tune aspects like gesture positioning, stating: *“everything is manually tunable, so I can position the boxes I can position the gestures.”* Presenters also commented on the quality of feedback from the gestural widgets. P5 found the experience to be better than expected come delivery time, pointing out that *“I expected it to be a bit more confusing and hard to handle, but it was a really good experience, I could easily communicate, understand why it’s not working or why it is working.”*

5.3.4 Challenges

In the course of evaluating the new presentation tool, both presenters and audience members highlighted certain challenges that they encountered. From the presenter’s perspective, the need for a more spontaneous interaction with the tool was evident. P6 expressed a desire for *“more dynamic control over charts,”* emphasizing that there should not be a necessity to *“pre-plan things”* extensively in advance. This suggests that presenters are seeking a more intuitive interface that allows for real-time adaptability during presentations, such as being able to go off-script, ad-lib or respond to audience interruptions and questions. Such feedback underscores the need for a balance between automation and manual control, ensuring that presenters can seamlessly navigate their material without feeling overly constrained by the tool. P3 suggested a feature to provide prompts or cues when specific gestures should be performed. This would serve as a helpful guide for the presenter, ensuring that their gestures align with the anticipated visual effects, thus improving the overall feedback loop during the presentation. On the other hand, the audience had concerns rooted in VisConductor’s animation and gestural coherence. For instance, a misalignment like a pointing gesture executed away from the actual animation could confuse the audience, highlighting the need for coherent gesture-animation interactions. A8 pointed out a potential disconnect between the gestures and the resulting visual effects, hinting at moments when gestures might *“not match the visual effects or properly communicate the message.”* Such feedback emphasizes the importance of intuitive gesture-to-visual mappings, ensuring that the audience can effortlessly follow along.

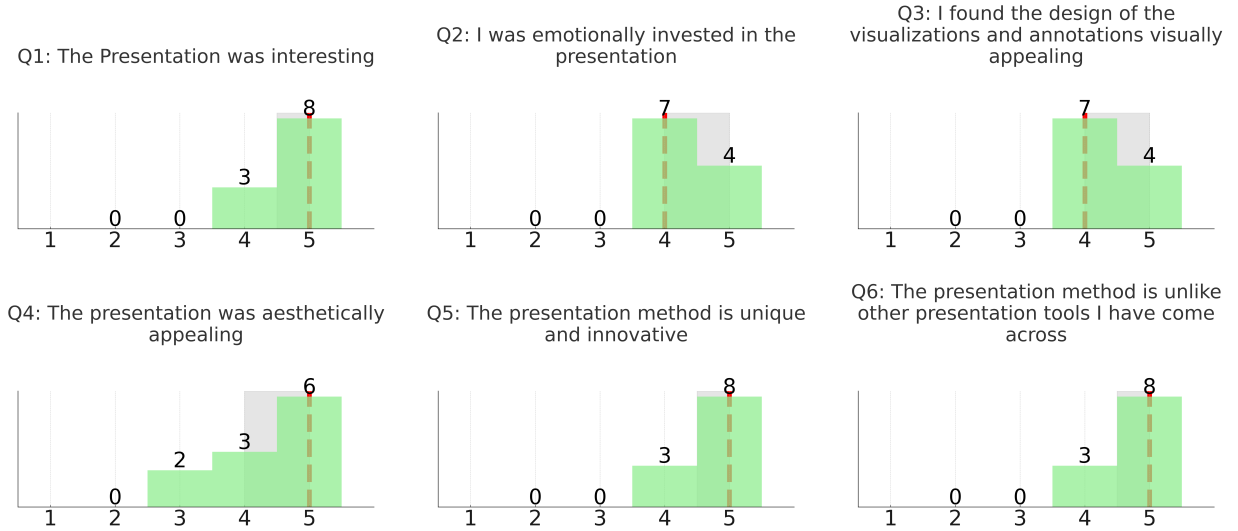


Figure 5.1: Distribution of responses on the engagement questionnaire of the audience study, where 1 = “strongly disagree” and 5 = “strongly agree”. The red dashed line indicates the median and the shaded region denotes the interquartile range (IQR).

5.4 Audience Engagement

As shown in Figure 5.1, the responses to the User Engagement Scale [50] questions (Q1-Q6) were gathered on a 5-point Likert scale, with the medians all being 4 and above. The predominant sentiment expressed by the respondents was overwhelmingly positive, with most emphasizing the unique and innovative nature of the presentation method. A significant majority found the presentations crafted with VisConductor to be engaging and distinct from other presentation tools they have encountered.

The interest level of the presentation (Q1) reveals a preponderance of positive feedback. Most audience members found the content and delivery of the presentation captivating. The dynamic animations and gesture-based interactions offered by VisConductor appear to play a vital role in sustaining the audience’s attention, transforming an otherwise static presentation into a riveting narrative. On emotional investment (Q2), the results indicate that the tool’s ability to blend animations with gesture-based interactions was appreciated. The audience found that these features, when executed well, intensified the emotional emphasis of a presentation, establishing a deeper connection between the presenter, the content, and the audience. When assessing the visual appeal of the design and annotations (Q3), the responses emphasize the effectiveness of VisConductor in transforming intricate

data patterns into compelling visual stories. Features such as visual foreshadowing and affective animations played a pivotal role in ensuring this visual appeal. For the presentation’s overall aesthetic quality (Q4), the ratings indicate that VisConductor is capable of helping presenters create presentations that strike a balance between data representation and aesthetic design, ensuring that the visual appeal complements the data narrative. The results for the novelty of the presentation method (Q5) suggest that for many, the experience of viewing a presentation crafted with VisConductor was refreshingly innovative, breaking away from conventional clicker-based narratives in remote settings as they separate content from the presenter. Lastly, in assessing the uniqueness of the presentation method compared to other tools (Q6), the overwhelmingly positive feedback is a testament to VisConductor’s potential to revolutionize remote data presentations.

Chapter 6

Discussion

Balancing flexibility and usability with configurable gesture widgets. We designed each gesture widget in VisConductor to be configurable, empowering presenters with the autonomy to fine-tune their interaction modalities, such as, the ability to adjust pose hold durations, reset pose detection post activation, or select which hand the widget should recognize. Yet, the real strength of VisConductor’s approach lies in its marriage of customizability with familiarity. It grounds advanced features in a familiar interface paradigm, ensuring that it is approachable by both novice and expert presenters. An alternative to the widget-based approach would be a system where gestures are detected indiscriminately, regardless of where they are performed within the webcam feed such as in Hall et al.[29], provided there is a visualization on screen. While this could offer broader flexibility, it might introduce unintended activations or reduced specificity in gesture recognition. The widget-based approach, in contrast, ensures a more tailored and precise interaction, with gestures being recognized in specific, user-defined regions.

In the realm of interactive presentations, ad-lib or unscripted formats stand distinct, bringing forth unique challenges when melded with gesture-based systems like VisConductor. These presentations, characterized by their spontaneity and fluidity, beckon for a tool that is not just reactive but proactively anticipatory, primed to intuit the presenter’s impromptu needs. At the heart of these challenges lies the adaptability of widgets, traditionally conceived for predetermined gestures. In the unpredictable landscape of ad-lib presentations, such as speakers responding to live audience questions, the need for “dynamic widgets” becomes evident. Importantly, their design should intuitively resonate with the spontaneous gesture required for an immediate operation. For example, if a presenter feels the impulse to underscore a particular data segment, the system could suggest a gesture widget tailored for highlighting. This on-the-fly widget recommendation can be

triggered based on the presenter’s verbal cues, akin to Liu et al.’s recommended visual captions [42]. Furthermore, this widget’s positioning should naturally align with the presenter’s motion, ensuring an organic interaction and the flawless execution of the intended operation.

Bridging functionality and expression with gestural interaction. Existing literature highlights a significant challenge faced when designing a system that recognizes and responds to operational gestures without conflicting with expressive gestures. This balance is pivotal, especially in live presentations where the presenter’s body language plays a crucial role in conveying information and building rapport with the audience. The feedback from our participants accentuated challenges like difficulty in precision pointing and the potential for unintentional triggering of actions. In VisConductor, the introduction of widgets with defined boundaries for operational gestures offers a potential solution to these challenges. These boundaries ensure that gestures are deliberate, allowing presenters to use other expressive gestures without interference. Such a design choice inherently respects the richness of human body language, ensuring that presenters can be expressive without the fear of unintentional system responses.

To further bolster this deliberateness, VisConductor is designed with distinct views for the presenter and the audience. The presenter’s view integrates feedback mechanisms, indicating how long poses should be held and when a specific gesture is triggered. Mirroring video game tutorials where players are guided on action duration or when a specific input is registered. This aids presenters in ensuring their gestures are deliberate and correctly recognized by the system. The audience view, on the other hand, offers a clean interface, devoid of these feedback elements, allowing for an immersive content experience. Further, the distinction between expressive and operational gestures, as discussed by Hall et al. [29], underscores the intricate challenges and vast potential of gesture-based systems. By allowing users to configure the duration a pose must be held before recognition or which hand the pose should be detected for, VisConductor adds an element of deliberateness to the interactions. Thus, fleeting expressive movements are less likely to be mistaken for operational gestures.

Providing feedback for gesture interaction in live presentations. The cognitive demands placed on presenters during live sessions cannot be ignored [56]. The anticipation and execution of the next move, combined with tracking the current state of the presentation, can be taxing. VisConductor addresses this by offering gesture widgets that provide real-time feedback, which transparently communicates the pose detection state. These widgets serve as visual cues, indicating how long poses should be held and when a specific gesture is triggered. The inclusion of labels further augments this feedback mechanism, allowing users to integrate metadata that can assist them during the presentation. Feed-

back’s role is not merely confirmatory. The widgets themselves can evolve into dynamic prompts, guiding the presenter through their narrative. Yet, while feedback is indispensable, it is essential to strike a balance. An overwhelming barrage of feedback cues might divert the presenter’s focus, detracting from the core content of their presentation.

Enhancing affective animations and emotional resonance with gesture integration. The power of visual storytelling in enhancing comprehension and retention has been well-documented, but VisConductor elevates this understanding by interweaving affective animations—that elicit or communicate emotions—with gesture integration. This union enhances the narrative’s emotional undertones. Diving into the significance of affective animations, their potential extends far beyond aesthetic appeal. As emphasized by A9, these animations can steer the audience’s focus, potentially even conveying emotional nuances, adding layers of depth to data narratives. Such capabilities are particularly crucial in fields like marketing or advocacy, where evoking emotions can substantially influence decision-making. For example, the dynamic unimanual *Dialling* gesture metaphorically symbolizes the progression of time. This gesture acts as a bridge between the narrative’s temporal aspects and the audience, allowing presenters to navigate through timelines. When combined with affective animations, especially those designed using ease functions and morphs, this gesture brings fluidity and a sense of continuity to the narrative. The eases and morphs, by their very nature, provide smooth transitions, ensuring that the data story flows seamlessly.

Chapter 7

Limitations and Future Work

Our evaluation showcased VisConductor’s potential in envisioning the next-generation augmented presentation authoring tools, but some limitations must be acknowledged. First, in our user study, presenters had only a short period to familiarize themselves with VisConductor, potentially preventing them from fully exploring the system’s capabilities. Future deployment studies with extended learning periods and allowing presenters to use their own content and speak ad-lib could offer a more comprehensive understanding of the tool’s capabilities.

Moreover, it is noteworthy that the current iteration of VisConductor is predominantly focused on animated scatter plots and bar chart races, which may not capture the complete spectrum of data visualizations that presenters might wish to utilize. Future development extending VisConductor’s concept to other types of animated visualizations (e.g., pie charts, geospatial maps, dynamic networks, geospatial origin/destination flow maps, alluvial diagrams etc.) is needed to fully explore this new design space of remote presentation authoring tools.

Third, while VisConductor has made strides in introducing affective animations, there is an opportunity to offer a more comprehensive toolset for customizing animations. An exciting direction could be developing a taxonomy for transitions between different animation states [30, 55]. This can empower presenters to select from pre-designed transition animations in literature to elevate their storytelling.

Lastly, some users’ feedback centered on the structured approach of VisConductor, while guaranteeing a consistent and controlled presentation flow, which highlights an opportunity to introduce more flexibility. Introducing gesture-driven *dynamic widgets* could

be a pivotal evolution. These widgets, specifically designed for on-the-fly initiation and positioning, would cater to the unpredictability of ad-lib presentations.

Chapter 8

Conclusion

We introduced VisConductor, an augmented video presentation tool that melds animated data visualizations with the innate power of hand gestures, which features a range of operations on visualizations including animation playback, visual foreshadowing, and annotation reveal. Our evaluation, encompassing both presenters and audiences, provided insights into the applicability and benefits of VisConductor. Presenters appreciated the enhanced control and dynamism, while audience members were captivated by the immersive and engaging nature of presentations crafted using VisConductor. As remote communication needs continue to grow, tools like VisConductor are poised to make data presentations not just a transfer of information, but a memorable and engaging experience.

References

- [1] Gesturing in the wild | Proceedings of the 24th BCS Interaction Specialist Group Conference.
- [2] Martha W. Alibali, Miriam Bassok, Karen Olseth Solomon, Sharon Syc, and Susan Goldin-Meadow. Illuminating mental representations through speech and gesture. *Psychological Science*, 10:327 – 333, 1999.
- [3] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Jason Leboe-McGowan, and Pourang Irani. Hooked on data videos: assessing the effect of animation and pictographs on viewer engagement. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, pages 1–9, Castiglione della Pescaia Grosseto Italy, May 2018. ACM.
- [4] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Jason Leboe-McGowan, and Pourang Irani. Hooked on data videos: assessing the effect of animation and pictographs on viewer engagement. *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, 2018.
- [5] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Andres Monroy-Hernandez, and Pourang Irani. Authoring Data-Driven Videos with DataClips. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):501–510, January 2017. Conference Name: IEEE Transactions on Visualization and Computer Graphics.
- [6] Apple Inc. Keynote. Apple Keynote, 2023. <https://www.apple.com/ca/keynote/>.
- [7] Narges Ashtari, Andrea Bunt, Joanna McGrenere, Michael Nebeling, and Parmit K. Chilana. Creating augmented and virtual reality applications: Current practices, challenges, and opportunities. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

- [8] Jose Maria Barrero, Nicholas Bloom, and Steven J Davis. The Evolution of Working from Home.
- [9] Thomas Baudel and Michel Beaudouin-Lafon. Charade: remote control of objects using free-hand gestures. *Communications of the ACM*, 36(7):28–35, July 1993.
- [10] BBC. Hans rosling’s 200 countries, 200 years, 4 minutes - the joy of stats - bbc four, 2012. YouTube video.
- [11] BBC. Hans rosling’s 200 countries, 200 years, 4 minutes - the joy of stats - bbc four, 2021. YouTube video.
- [12] Bostock, Michael and Ogievetsky, Vadim and Heer, Jeffrey. D3.js. D3.js, 2023. <https://d3js.org/>.
- [13] Matthew Brehmer, Bongshin Lee, Petra Isenberg, and Eun Kyoung Choe. A Comparative Evaluation of Animation and Small Multiples for Trend Visualization on Mobile Phones. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):364–374, January 2020. arXiv:1907.03919 [cs].
- [14] Jana Bressemer and Silva H. Ladewig. Rethinking gesture phases: Articulatory features of gestural movement? 2011.
- [15] Canva Pty Ltd. Canva. Canva, 2023. <https://www.canva.com/>.
- [16] Xiang Cao, Eyal Ofek, and David Vronay. Evaluation of alternative presentation control techniques. In *CHI ’05 Extended Abstracts on Human Factors in Computing Systems*, pages 1248–1251, Portland OR USA, April 2005. ACM.
- [17] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, May 2019. arXiv:1812.08008 [cs].
- [18] Adelia Carstens. Advice on the use of gestures in presentation skills manuals: alignment between theory, research and instruction. *Image & Text*, 2019.
- [19] Fanny Chevalier, Nathalie Henry Riche, Catherine Plaisant, Amira Chalbi, and Christophe Hurter. Animations 25 years later: New roles and opportunities. *Proceedings of the International Working Conference on Advanced Visual Interfaces*, 2016.
- [20] Zoom Video Communications. Zoom meetings, 2021. Video conferencing software.

- [21] Stefania Cuccurullo, Rita Francese, Sharefa Murad, Ignazio Passero, and Maurizio Tucci. A gestural approach to presentation exploiting motion capture metaphors. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 148–155, Capri Island Italy, May 2012. ACM.
- [22] Jean-Daniel Fekete, Jarke J. van Wijk, John T. Stasko, and Chris North. The Value of Information Visualization. In Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North, editors, *Information Visualization: Human-Centered Issues and Perspectives*, Lecture Notes in Computer Science, pages 1–18. Springer, Berlin, Heidelberg, 2008.
- [23] Danyel Fisher. Animation for Visualization: Opportunities and Drawbacks.
- [24] Adam Fourney, Michael Terry, and Richard Mann. Gesturing in the Wild: Understanding the Effects and Implications of Gesture-Based Interaction for Dynamic Presentations. 2010.
- [25] Gapminder Foundation. Gapminer. GapMiner, 2023. <https://www.gapminder.org/>.
- [26] S. Goldin-Meadow. *Thinking with Your Hands: The Surprising Science Behind How Gestures Shape Our Thoughts*. Basic Books, 2023.
- [27] Inc. GreenSock. Gsap: The standard for javascript html5 animation, 2023. Version 3.
- [28] Thomas C. Gunter, J. E. Douglas Weinbrenner, and Henning Holle. Inconsistent use of gesture space during abstract pointing impairs language comprehension. *Frontiers in Psychology*, 6, 2015.
- [29] Brian D. Hall, L. R. Bartram, and Matthew Brehmer. Augmented chironomia for presenting data to remote audiences. *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, 2022.
- [30] Jeffrey Heer and Ben Shneiderman. Interactive Dynamics for Visual Analysis.
- [31] Apple Inc. Use reactions, presenter overlay, and other effects when videoconferencing on mac, 2023.
- [32] Seokmin Kang and Barbara Tversky. From hands to minds: Gestures promote understanding. *Cognitive Research: Principles and Implications*, 1, 2016.
- [33] Adam Kendon. *Gesture Visible Action as Utterance*. Cambridge University Press, 1 edition, September 2004.

- [34] Robert Kosara and Jock Mackinlay. Storytelling: The Next Step for Visualization. *Computer*, 46(5):44–50, May 2013.
- [35] Praveen Kumar, Anurag Jaiswal, B.B.V.L. Deepak, and G. Ram Mohana Reddy. Hand gesture-based stable powerpoint presentation using kinect. 2018.
- [36] Google Creative Lab. Teachable machine, 2021. A web-based tool that makes creating machine learning models fast, easy, and accessible to everyone.
- [37] Xingyu Lan, Yang Shi, Yanqiu Wu, Xiaohan Jiao, and Nan Cao. Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics*, PP:1–1, 2021.
- [38] Elsie Lee-Robbins and Eytan Adar. Affective Learning Objectives for Communicative Visualizations, August 2022. arXiv:2208.04078 [cs].
- [39] Wenchao Li, Yun Wang, Haidong Zhang, and Huamin Qu. Improving engagement of animated visualization with visual foreshadowing. *2020 IEEE Visualization Conference (VIS)*, pages 141–145, 2020.
- [40] Jian Liao, Adnan Karim, S. Jadon, Rubaiat Habib Kazi, and Ryo Suzuki. Realitytalk: Real-time speech-driven augmented presentation for ar live storytelling. *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, 2022.
- [41] Xingyu Bruce Liu, Vladimir Kirilyuk, Xiuxiu Yuan, Alex Olwal, Peggy Chi, Xiang "Anthony" Chen, and Ruofei Du. Visual captions: Augmenting verbal communication with on-the-fly visuals. In *Proc. ACM CHI*, 2023.
- [42] Xingyu "Bruce" Liu, Vladimir Kirilyuk, Xiuxiu Yuan, Alex Olwal, Peggy Chi, Xiang "Anthony" Chen, and Ruofei Du. Visual Captions: Augmenting Verbal Communication with On-the-fly Visuals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–20, New York, NY, USA, April 2023. Association for Computing Machinery.
- [43] Microsoft Corp. Microsoft excel. Microsoft Excel, 2023. <https://www.microsoft.com/en-us/microsoft-365/excel>.
- [44] Microsoft Corp. Microsoft power bi. Microsoft Power BI, 2023. <https://powerbi.microsoft.com/>.
- [45] Microsoft Corp. Microsoft powerpoint. Microsoft PowerPoint, 2023. <https://www.microsoft.com/en-ca/microsoft-365/powerpoint>.

- [46] Michael Nebeling. Rapid prototyping of augmented reality & virtual reality interfaces. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.
- [47] Michael Nebeling, Janet Nebeling, Ao Yu, and Rob Rumble. Protoar: Rapid physical-digital prototyping of mobile augmented reality applications. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.
- [48] Michael Nebeling and Maximilian Speicher. The trouble with augmented reality/virtual reality authoring tools. *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 333–337, 2018.
- [49] Heather L. O’Brien and Elaine G. Toms. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6):938–955, April 2008.
- [50] Heather L. O’Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112:28–39, April 2018.
- [51] Pedersen, Thomas Lin. gganimate. gganimate, 2023. <https://gganimate.com/>.
- [52] Plotly Technologies Inc. Plotly. Plotly, 2023. <https://plotly.com/>.
- [53] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyoung Choe. Chartaccent: Annotation for data-driven storytelling. *2017 IEEE Pacific Visualization Symposium (PacificVis)*, pages 230–239, 2017.
- [54] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, Sheelagh Carpendale, Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. *Data-Driven Storytelling*. A K Peters/CRC Press, 1 edition, March 2018.
- [55] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko. Effectiveness of Animation in Trend Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1325–1332, November 2008.
- [56] Nazmus Saquib, Rubaiat Habib Kazi, Li yi Wei, and Wilmot Li. Interactive body-driven graphics for augmented video performance. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.
- [57] Andreas Schallwig. Fingerpose, 2023. 0.1.0.

- [58] E Segel and J Heer. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, November 2010.
- [59] Minjeong Shin, Joohee Kim, Yunha Han, Lexing Xie, Mitchell Whitelaw, Bum Chul Kwon, Sungahn Ko, and Niklas Elmqvist. Roslingifier: Semi-automated storytelling for animated scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 29:2980–2995, 2022.
- [60] Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1):116–124, 2013.
- [61] Maximilian Speicher and Michael Nebeling. Gesturewiz: A human-powered gesture design environment for user interface prototypes. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.
- [62] J. Steele and N. Iliinsky. *Beautiful Visualization: Looking at Data through the Eyes of Experts*. Theory in practice series. O’Reilly Media, 2010.
- [63] Charles D Stolper, Bongshin Lee, Nathalie Henry Riche, and John Stasko. Emerging and Recurring Data-Driven Storytelling Techniques: Analysis of a Curated Collection of Recent Stories.
- [64] Charles D Stolper, Bongshin Lee, Nathalie Henry Riche, and John Stasko. Emerging and Recurring Data-Driven Storytelling Techniques: Analysis of a Curated Collection of Recent Stories.
- [65] Ryo Suzuki, Rubaiat Habib Kazi, Li yi Wei, Stephen DiVerdi, Wilmot Li, and Daniel Leithinger. Realitysketch: Embedding responsive graphics and visualizations in ar with dynamic sketching. *Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 2020.
- [66] Cisco Systems. Cisco webex, 2021. Video conferencing, online meetings, and web conferencing software.
- [67] Tableau Software. Tableau. Tableau Software, 2023. <https://www.tableau.com/>.
- [68] Mmhmm Team. mmhmm, 2021. Interactive video presentation software.
- [69] Runway Team. Runwayml, 2021. Machine learning for creators.

- [70] John Thompson, Zhicheng Liu, and John Stasko. Data Animator: Authoring Expressive Animated Data Graphics. In *CHI '21: Proceedings of the ACM Conference on Human Factors in Computing Systems*, May 2021.
- [71] Lyn Tieu, Robert Pasternak, Philippe Schlenker, and Emmanuel Chemla. Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa*, 2, 2017.
- [72] Julien Valentin, Adarsh Kowdle, Suhil Nassar, Ameesh Makadia, Angjoo Kanazawa, Michael Ratasich, and Caroline Pantofaru. Mediapipe: A framework for building perception pipelines. *Google Research*, 2019.
- [73] Petra Wagner, Zofia Malisz, and Stefan Kopp. Gesture and speech in interaction: An overview. *Speech Commun.*, 57:209–232, 2014.
- [74] Longqi Yang, David Holtz, Sonia Jaffe, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce, Neha Parikh Shah, Kevin Sherman, Brent Hecht, and Jaime Teevan. The effects of remote work on collaboration among information workers. *Nature Human Behaviour*, September 2021.
- [75] Fan Yujie, Xu Hui, and Jiang Jie. A survey of time series data visualization research. 2020.
- [76] Haipeng Zeng, Xingbo Wang, Yong Wang, Aoyu Wu, Ting-Chuen Pong, and Huamin Qu. Gesturelens: Visual analysis of gestures in presentation videos. *IEEE Transactions on Visualization and Computer Graphics*, PP:1–1, 2022.