Dais: A Live Presentation and Analytics Tool Utilizing Google Glass

Elaine Zhou
Stanford University
450 Serra Mall
Stanford, CA 94305 USA
ezhou@stanford.edu

Harrison Wray
Stanford University
450 Serra Mall
Stanford, CA 94305 USA
hwray@stanford.edu

Brian Lam
Stanford University
450 Serra Mall
Stanford, CA 94305 USA
brianlam@stanford.edu

Adrian Rodriguez
Stanford University
450 Serra Mall
Stanford, CA 94305 USA
Arod2016@stanford.edu

Abstract
This report details Dais: a live presentation aid and analytics tool for the Google Glass and web. Our team's aim is to help oral presenters improve in two ways: (1) displaying live feedback indicating quality of visual spread and speaking volume, and (2) offering a post-presentation analytics and visualization platform for this data. Dais improves on existing presentation aid tools by leveraging the inherent strengths of wearable devices: the ability to analyze a constant stream of real-world sensor data to interpret and classify user behaviors. The ultimate goal of Dais is to reinforce good presentation techniques and to turn these behaviors into intuitive habits, eliminating the need for users to rely on Dais or other external aids while presenting.

Author Keywords
Wearable Computing, Presentation Aid, Human Factors

ACM Classification Keywords
H.5.1. Information interfaces and presentation: Multimedia Information Systems

Introduction
Psychologists have extensively studied the impact of Presentations are a crucial form of communicating ideas in the modern world. However, many modern presentation aids serve as a crutch at best and a

distraction at worst. They often focus on matters of convenience: making the presenter's notes more accessible, making slide changes easier, etc. Most presentation aids ignore the fundamental elements that characterize a good presentation: balanced visual spread, consistent speaking volume, etc. Here, we discuss and consider a subset of related work on methods for improving public speaking.

Gaze and Visual Spread

Kleinke, University of Alaska psychologist and leading author on eye gaze, concludes that when speakers establish eye contact, they are perceived as more credible, sociable, likable, competent, and attentive by others [1]. Thus, a speaker's visual spread is a vital metric for determining a presentation's effectiveness. Dais determines whether visual spread and effective speaking are correlated by combining quantitative visual spread measurements with audience perception. Kleinke's studies determine eye contact by having human evaluators manually record every time they believe two subjects have made eye contact. With the use of digital sensors, we believe that Google Glass will yield more accurate determinations of whether visual spread and presentation effectiveness are correlated.

Public Speaking Anxiety Virtual Reality Therapy

Anderson, et al., observed significant public speaking improvement in every subject they treated with virtual reality therapy [2]. Speaking phobic subjects were exposed to several virtual speaking scenarios through head-mounted virtual reality displays, and spoke to large, disruptive audiences. After the treatment, every subject self-reported decreased speaking anxiety, thus suggesting that virtual reality speaking exercises may help people improve their speaking ability [2]. This assumption inspired Dais's implementation of live presentation feedback. Since Google Glass doesn't require much hardware, we believe that our feedback implementation is more practical. Moreover, since our feedback is algorithmically generated-- which this study's was not-- we believe that Dais's feedback can be more productive.

Palette: Visual Aid Interactions

In the past decade, many attempts have been made to Palette is an interface in which presentation content is represented by encoded cards inspired by ordinary, physical note cards intended to improve public speaking by enhancing slideshow interactions [3]. As a user holds up a card, its corresponding content is projected on screen. The ease in perceiving these cards focally and peripherally enhanced collaborative presentation as well as spontaneous and brief discussions. While Palette demonstrated an ability to make visual aid interactions more seamless, this benefit was shown to scale poorly. Almost every presenter in this study operated several presentation props at once. A later re-evaluation of the interface found the system to have too many complex interdependencies to fluidly support presenters' tasks [3]. Therefore, we decided against including visual aid interaction from our presentation system.

Ultimately, public speaking is most effective when natural and uncluttered. From our literature review, we concluded Google Glass to be the least intrusive and potentially most effective presentation aid.

Dais System Design Overview

To the user, Dais is composed of two components: a Google Glass application to provide live feedback and a web component to show post-presentation analytics. Here we describe our user-centered design methodology involving needfinding, prototyping, as well as the technical and design decisions and rationales that our team made in these two main components.

Needfinding

We observed and talked with novice speakers and interviewed two professional oral communicator lecturers -- Matt Vassar, in the Technical Communication Program, and Sohui Lee, PhD., Assistant Director of the Hume Writing Center. In these conversations, we learned one of the biggest problems students have is maintaining steady eye gaze and vocal variation. Video recording, the most common method, is typically ineffective due to its awkwardness. Conversations and observations of novice speakers surfaced further issues. Psychologically, "time and sound [are] so warped when you're presenting." When told to pay attention to a particular presentation aspect, students start off with a strong performance, but then "[forget] to pay attention to that." Immediate feedback is useful, but only in small doses. From this exploration, narrowed our problem space: non-professional oral presenters need a way to receive small amounts of immediate feedback and then later analyze more details of their performances.

Live Feedback – Google Glass

The main feature we wanted to prototype, build, and test was live feedback microinteractions for eye gaze and volume. In order to fully embrace wearable paradigms, this live feedback should be less than 4 seconds, to avoid distracting and interrupting the user during their presentation.

Live feedback is determined relative to several baseline calibration readings, which are taken before the presentation. During the calibration phase, the user looks to the left and right sides of the room, and taps to record compass headings for those orientations using the Glass magnetometer. The user also taps to record several seconds of background noise from the microphone, to establish a noise floor for the room; and lastly taps to record the first several seconds of their speech, to establish an ideal voice volume.

By continually reading sensor data and comparing to these baseline calibrations, live feedback can be offered according to several heuristic thresholds. If the user looks far outside the left and right sides of the room, Dais assumes the user has turned their head away from the audience to read off their own slides (a classic presentation blunder), and displays text feedback telling the user to "Face forward!" If the up-down pitch of the user's gaze falls below a certain negative threshold, Dais assumes the user is looking downwards to read from a script, and quickly prompts the user to look back up at the audience. If the user's gaze becomes fixed on the left or right side of the room for a significant time, Dais offers directional feedback telling the user to increase visual spread by looking at the other half of the audience.

Lastly, if the Glass microphone detects several seconds of low-volume (but above-floor) audio input, it assumes the user is mumbling, and prompts the user to "Speak up!" This feature was a result of needfinding, when students noted that they had the biggest problem staying engaged when they couldn't hear the speaker.

Post-Presentation Analytics – Web App

Taking these results into account, we built a real-time app platform on which the presenter can view important metrics from any previously given presentation. The presentation information gathered from the sensors (head tilt and position from the accelerometer, volume from the microphone) are packaged and sent to the online application. Via this platform, we display a heat map of head position to model visual spread and a line graph charting the progression of the speaker's volume. Since the goal of post-presentation evaluation is pinpointing and correcting errors made during the presentation, we designed the analytics with the intention of highlighting errors made. With this paradigm in mind, the heat map highlights the user's tendency to stare at any single point, and the volume line graph indicates the user's volume relative to an inputted "desired speaking level."

Evaluation and Discussion

Our primary area of evaluation was how to minimize the level of cognitive and visual disruption incurred by the live feedback cues. A key interface decision was whether to prompt users with spatial cues (light up the left side of the screen) or directional cues (show an arrow pointing to the left). Marotta, et al. addressed two different modes of attentional selection – gaze versus arrows as spatial cues. They found that attention is nonspecifically directed to nearby objects when an arrow is used as a guiding cue, whereas it is selectively directed to a specific cued location when eye gaze is used [4]. Given this insight, we prototyped these two methods, and general heuristic feedback followed Marotta's hypothesis (arrows over spatial cues). Additionally, arrows with heavier ends (→ versus >) were easier to glance at, thus better on Google Glass.

Future Work

The initial implementation of Dais barely scratches the surface of the analytic feedback that can be achieved through Glass's sensors. The accelerometer could be used as a pedometer to detect footstep patterns, allowing Dais to prompt overly-stationary presenters to "Move more," or tell nervous presenters to "Stand still" or "Stop pacing." Glass' background speech recognition could count the number of times a presenter says "uh" or "um," or give live feedback telling them to "Speed up" or "Slow down" their speech. Head motion can potentially correlate with artifacts from shoulder and arm gestures, enabling Dais to give feedback on users' body language using only accelerometer data from the head. Ultimately, we'd like to extrapolate movement, voice, and gaze analytics to form an advanced database of presentation profiles, allowing users to view a summary of individual strengths and weaknesses, and different presentations styles based on professionals.

Conclusion

In the course of our research, we designed and implemented a live presentation aid and analytics tool, which focuses on promoting balanced eye gaze spread and consistent voice volume. We conclude that Dais in its current form has the potential to improve the skills of non-professional speakers preparing for a presentation. With further research and development into some of the work mentioned above -- intelligent modeling algorithms, footstep and gesture recognition, detailed presentation profiles -- we believe that Dais could offer even more useful feedback and analytics for users. In this team's vision, a full and ideal implementation of Dais will train users to adopt good presentation habits intuitively, eliminating their need to rely on Dais or other aids for future presentations.

References

[1] Kleinke, Chris L. "Gaze and eye contact: a research review." *Psychological bulletin* 100.1 (1986): 78.

[2] Anderson, P. L., Zimand, E., Hodges, L. F., & Rothbaum, B. O. (2005). Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and anxiety*, *22*(3), 156-158.

[3] Nelson, L., Ichimura, S., Pedersen, E.R. and Adams, L. Palette: A Paper Interface for Giving Presentations, in Proceedings of CHI 99, ACM Press, pp. 354-361, 1999.

[4] Marotta, A., Lupiáñez, J., Martella, D. and Casagrande, M (2012). Eye gaze vs. arrows as spatial cues: Two qualitatively different modes of attentional selection. *Journal of Experimental Psychology*