Funky Wall: Presenting Mood Boards Using Gesture, Speech and Visuals

Andrés Lucero
Department of Industrial Design
Eindhoven University of Technology
5600MB Eindhoven, the Netherlands
a.a.lucero@tue.nl

Dzmitry Aliakseyeu
Media Interaction Group
Philips Research Eindhoven
5656AA Eindhoven, the Netherlands
dzmitry.aliakseyeu@philips.com

Jean-Bernard Martens
Department of Industrial Design
Eindhoven University of Technology
5600MB Eindhoven, the Netherlands
j.b.o.s.martens@tue.nl

Figure 1. Presenting the story behind a mood board using gestures, speech and visuals.

ABSTRACT
In our studies aimed at understanding design practice we have identified the creation of mood boards as a relevant task for designers. In this paper we introduce an interactive wall-mounted display system that supports the presentation of mood boards. The system allows designers to easily record their mood board presentations while capturing the richness of their individual presentation skills and style. Designers and clients can play back, explore and comment on different aspects of the presentation using an intuitive and flexible interaction based on hand gestures thus supporting two-way communication. The system records the presentation and organizes it into three information layers (i.e. gesture, sound and visuals), which are first used to segment the presentation into meaningful parts, and later for playback. Exploratory evaluations show that designers are able to use the system with no prior training, and see a practical use of the proposed system in their design studios.

Categories and Subject Descriptors
H.5.m [**Information Interfaces & Presentation**]: Miscellaneous.

General Terms
Design, Human Factors, Performance.

Keywords
Gesture-based interaction, wall projection displays.

AVI'08, May 28-30, 2008, Napoli, Italy

1. INTRODUCTION
Designers commonly use mood boards in the early stages of the design process [4], to explore, communicate, and discuss ideas together with their clients. These boards can be created with different types of media although designers usually use images to say something about the target audience, product, and/or company they are designing for. What may be easily overlooked, however, is that there is a story behind every mood board.

Once the mood board is completed, designers must communicate the story (and the ideas) behind the mood board. Usually designers will meet their clients to directly present, discuss and receive feedback on their mood boards. However, in large companies mood boards are uploaded to the company's Intranet so they can be experienced by and inspire different departments (e.g. design, marketing, sales, etc.). It is also common that clients and the design team itself are distributed over the globe, working in different time zones. Mood boards are then embedded in PowerPoint presentations and attached to an extra A4 text document that explains the mood board. In these cases, Intranet or PowerPoint presentation, the main question is, how can designers make sure that the right message is conveyed? Why was a given image chosen? What is the path through the mood board that the designer intended in order to tell the story? And equally important, how can clients reply and give feedback on what they are thinking? More generally speaking, how can we support presenting and receiving feedback for a mood board?

We propose an interactive system called 'Funky Wall' (Figure 1) that supports the presentation of mood boards by recording and keeping essential aspects of the presentation at three main information layers: gesture, sound (speech) and visuals. These information layers are first analyzed in order to segment the presentation in a meaningful way. Each segment is then associated with a specific time, interval and area on the mood board allowing designers and clients to experience the mood board, taking into account both time and space.

2. BACKGROUND

The field of human-computer interaction (HCI) has been investigating how people interact with computer systems at work (and more recently at home), trying to help them achieve their goals. Within HCI, researchers have already identified the potential behind interactive vertical surfaces as a more natural and familiar setting to address design (collaborative) interactions. The ID-MIX project [9] tries to assess the relevance and impact of augmented reality systems in design practices. The question the project addresses is if professional users (i.e. industrial designers) are willing to change their current work practices when confronted with alternative 'augmented reality' approaches. To gain a better understanding of design practice, we have conducted several user studies using diverse methods. By using probes in a professional context [10] we were able to identify a relevant task for designers: making mood boards. Subsequently, we have conducted contextual inquiries with Dutch industrial designers and interviews with Finnish fashion and textile designers to get a better understanding of why designers use mood boards and how they create them. In these studies, we have identified five stages of the mood-board making process: 1) 'collecting', 2) 'browsing', 3) 'piling', 4) 'building', and 5) 'presenting'.

3. RELATED WORK

Clark and Brenan [2] have extensively studied the relation between gestures and speech, and the role of gestures in human communication. Clark and Brenan argue that gestures together with communicative statements help establish common understanding, and that an appropriate gesture that is easily interpretable is preferable over complex sentence constructions. Gestures have also been widely explored as a natural way of interaction for a range of systems such as tabletop, vertical displays, multi-device environments, and 3D virtual environments. An example of a public display system that is controlled by gestures was presented in [13]. The authors aimed at studying shared, interactive public displays that support transition from implicit to explicit interaction. They used hand gestures and touch for explicit interaction, while body orientation and location played part in implicit interaction. A few systems employed gesture-based interaction in addition to speech, for either enriching the presentation process or to improve the communication with remote parties. The Charade system [1] allows presenters to use free-hand gestures to control a remote computer display, while also using gestures for communicating with the audience. Kirk et al. [7] studied different ways to represent gesture shadows (hands, hands and sketch, sketch only). They concluded that unmediated video representations of hands speed up performance without affecting accuracy. There is also a large area of research that looks at optimal meeting content capturing and browsing [5]. Many of these systems are based on the idea of Activity-based Information Retrieval, which proposes to use user activity (such as note-taking, annotating, writing on whiteboards) to index multimedia data and make data retrieval easier [8]. However only a few examples can be found where a speech plus gesture approach is used to enrich the capturing and (re)viewing of presentations. Ju et al. [6] use a motion estimation technique to detect key frames and segment the video (recorded presentation). Another example is the Active Multimodal Presentations [3] concept. The main difference with the Funky wall is that it addresses offline communication and attempts to create a structure using only implicit information (speech and gestures) for segmentation.

4. DESIGNING THE 'FUNKY WALL'

From the five stages of the mood-board making process, we have conducted exploratory studies in relation to 'browsing' [9]. We now focus our work on supporting the final stage, 'presenting', by designing a 'Funky Wall' that: 1) allows designers to easily record their mood board presentations while capturing the richness of their individual presentation skills and style, 2) allows both designers and clients to play back and explore different aspects of the presentation using an intuitive and flexible interaction involving hand gestures, and 3) supports two-way communication needed for successful mood-board design, by allowing clients to reply and share their thoughts on the mood board contents.

4.1 Proximity-Based Interaction

The 'Funky Wall' employs four different ranges of interaction (Figure 2) depending on the designer's proximity to the mood board: 'presenting', 'contemplating', 'replaying', and 'exploring'. Different functionalities are made available for each range. Gesturing close to the screen is used to record a presentation (<0.5m). When the presentation has been created, designers or clients can then 'contemplate' the mood board from a distance (>2m, no gestures), they can 'replay' the entire presentation (gesturing 1.5-2.m), or they can also 'explore' specific parts of the recorded presentation (gesturing 0.5-1.5m). Our four ranges of interaction resemble the ranges proposed in [13], and [11].

4.2 Intuitive & Flexible: Gestures & Speech

From our studies we have learned that for activities involving creation designers prefer working with their hands and with tools that allow flexibility and intuitive interaction (e.g. pencil and paper). To keep the interaction simple, designers can record their presentation by gesturing and explaining the mood board in front of the screen, using their hands to point or outline specific areas of the mood board. Preliminary observations show that location and speed of the gesture can be used to create meaningful indices, i.e., to associate the speech layer with a particular area of interest.

4.3 Two-Way Communication

A mood board is an idea development tool. During the mood-board making process, designers and clients have several rounds of discussions to reach agreement on the ideas being presented in the mood board. Therefore, for a successful mood-board design the tool should support two-way communication between designer and client. The 'Funky Wall' supports this iterative process by allowing designers and clients to provide input by creating a presentation and share their thoughts by providing feedback. For this type of communication to happen, two 'Funky Walls' are needed, one for the designer and another for the client.

5. INTERACTION TECHNIQUES

5.1 Presenting

To begin recording their presentation, designers simply need to gesture and speak next to the screen (<0.5m) (Figure 2a). The system displays white traces of the gestures made, as if designers were putting down a continuous flow of paint with their hands. To allow good visibility of the mood board the opacity of the white trace is set to 30%. The system captures and segments the speech and the natural hand movements made by the designer, creating associations between audio segments and gestures.

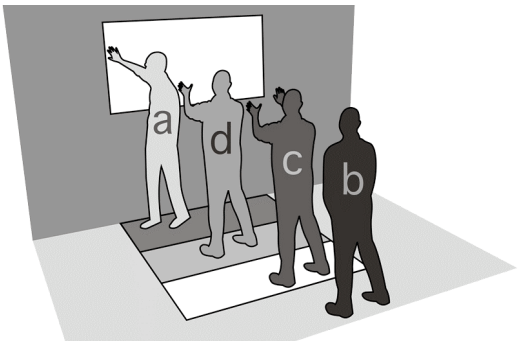

Figure 2. Interaction modalities revealed based on proximity. (a) 'Presenting' by gesturing next to the screen (<0.5m), (b) 'contemplating' the mood board (no gestures >2m), (c) 'replaying' the entire presentation (gesturing 1.5-2m), and (d) 'exploring' parts of the presentation (gesturing 0.5-1.5m).

5.2 Contemplating

Once a presentation has been completed, designers or clients can contemplate the mood board from a distance (>2m) (Figure 2b) for a comfortable overview. No gesturing is possible at this range.

5.3 Replaying

Spectators can replay the entire presentation by approaching the screen (between 1.5-2m) (Figure 2c). Raising the dominant hand results in displaying all gestures made during the presentation semitransparent on top of the mood board. Raising the non-dominant hand will trigger the complete recorded speech or audio explanation. By putting both hands together, both the recorded speech and the dynamic gestures unfold as the presentation progresses. Having an overview of all gestures allows spectators to quickly see areas of high interest where gestures concentrate.

5.4 Exploring

Taking one step closer towards the screen allows exploring specific parts of the presentation (between 0.5-1.5m) (Figure 2d). By pointing with the dominant hand to a given area in the mood board, users can view a static representation of the traces made in that area. These overlaid traces of gestures serve as guides for retrieval. Putting both hands together will display the dynamic gestures together with its corresponding spoken explanation. The mood board remains visible throughout the exploration process. To provide visual contextual feedback within the presentation, the tool highlights the explanations made by the designer just before and immediately after the current gesture. The previous gesture is shown in a lighter shade of white as if faded. The next gesture is displayed in black, as something that still needs to be discovered. Mood-board presentations last somewhere between 5 and 8 minutes. Therefore if the designer is unsatisfied with the results of the presentation, we propose that they present once again instead of providing a tool that allows editing specific parts.

5.5 Supporting Two-Way Communication

To truly support two-way communication, clients must be able to give designers feedback based on their perception and interpretation of the mood board. By having a similar 'Funky Wall' in their office, clients can explore the entire presentation (or parts of it), and later reply by adding their own comments to the mood board using the same interaction modalities described in 'presenting', 'contemplating', 'replaying', and 'exploring'.

6. EVALUATION

We conducted an exploratory user study of the 'Funky Wall' to test its usefulness and usability. First, we wanted to see if practicing designers would see the prototype as a relevant tool to present their mood boards. Second, we wanted to test the interaction techniques in terms of naturalness, ease of learning and use. We recruited five practicing designers with at least 5 years of experience. The participants varied in gender (1 female, 4 male), age (between 30 and 40), and preferred hand (4 right-handed, 1 left-handed). The evaluations were conducted individually.

6.1 Tasks

In the first part of the study participants created their own story for a mood board we gave them (approx. 5 minutes). Each participant was told that they would be using a system that tracked and displayed traces of their hand movements. In the second part participants explored an existing presentation using the system. Following a brief description of the interaction we allowed them to freely explore the functionality and get acquainted with the system (approx. 10 minutes). In the third part we asked them to walk us through their experience using the system (approx. 30 minutes per participant). All sessions were recorded on video.

6.2 Implementation

The system was set up using a desktop PC connected to a back-projection screen of size 2.0x1.5m (1024x768 pixels), as well as an ultrasonic tracking system – InterSense IS-600 used to track hands. During the sessions participants wore custom-designed interaction gloves that contained the sensors. The application was written in C# and used OpenGL for visualization purposes. The presentation and replay parts were fully functional. The analysis phase, where the presentation is segmented, was done manually.

6.3 Findings

6.3.1 Participants Agreed on the Principles

Designers were positive about the general underlying principles of the system to support the presentation of mood boards:

"This system helps you get the explanation the designer intended. I can experience the thoughts behind the images. Clients will have their own associations and thoughts behind these images." [P1]

"In case of long and complex presentations this allows you to have reminders of where certain parts were, like chapters. You see the entire mood board and you can zoom into parts." [P3]

6.3.2 Hand Gestures

In the first part of the study, designers were able to interact with the system with no prior training, especially liking the naturalness and simplicity of interaction. However, in the second part designers began to experience some difficulties when exploring:

"Bringing both hands together to trigger sounds is very uncomfortable. Maybe a quick movement in the air to press." [P4]

"I found it a bit difficult to navigate to the next and the previous thing. Maybe a flick of the wrist in a given direction to the sides should allow you to go forward and backwards." [P5]

6.3.3 Visual Feedback

Regarding the visual feedback provided by the system (i.e. traces of gestures on top of the mood board), designers first reflected on the amount of visual clutter, and had different opinions:

"At a certain point it is getting increasingly cluttered." [P4]

"The way the visual feedback is presented is done in a subtle way; it does not ruin the impression of the mood board." [P5]

Participants also commented on the helpfulness of playing back gestures dynamically as they heard the explanation:

"It helps to better explain the picture. It gives a touch of sensibility. It makes it easier to connect. Although you are not present, it seems that you are there. It is like a ghost of you." [P1]

"It really (makes it) much more alive. I can feel that the designer was there doing those gestures. It makes it more human." [P2]

Finally, designers reflected on the usefulness of having contextual feedback for the current, previous, and next speech segment:

"In traditional presentations you have no cues about what is happening. This is much more intuitive than just having a timeline or something similar because now you can actually see how things unfold temporally alongside the thematic unfolding." [P5]

7. DISCUSSION

7.1 Feasibility of the System

In our prototype the analysis phase, where the presentation is segmented, was done manually. The main reason for doing this was that the goal of the study was to first assess the potential usefulness and usability of such a system. However, based on the results reported in the literature and the analysis of gesture-speech synchronization automation, our system seems feasible [12].

For segmentation our system does not need to recognize speech, we only need to detect phrase boundaries. One way of detecting phrase boundaries is by using pauses (intervals of non-speech audio between speech segments) [14]. Stifelman [12] found that phrases could be robustly identified using a threshold of 155 ms; pauses shorter than the threshold are most likely pauses within a phrase while longer ones are pauses between phrases. The speed and location of gestures can also be used to make the segmentation more robust. In our study we observed that speed could be used to separate between explanations of specific parts (slow movements), connections between different parts (fast long movements), and the general discussion of the mood board (often fast short movements).

7.2 Using Other Media to Record and Replay

We believe the use of gestures enriches the presentation of mood boards by allowing designers to clearly express the feelings and the ideas contained. The same is applicable to the replaying and annotating of the presentation. However the latter part can also be done on any desktop system using a standard pointing device such as a mouse. In principle, the presentation could also be done on a desktop but we fear that the added richness will be lost.

8. CONCLUSION

We built and evaluated a system that supports designers in conveying the story behind mood boards in situations when face-to-face communication is not possible. The 'Funky Wall' allows designers to easily record mood board presentations while capturing the richness of their individual presentation skills and style. It also allows both designers and clients to play back, explore and comment on different aspects of the presentation using an intuitive and flexible gesture-based interaction. We have evaluated the system with professional designers in order to test its usefulness and usability. The results of the study showed that designers saw a practical use of the system in their design studios. Participants felt that the system gave them control over the presentation, so they could, with little effort, explore different aspects of the mood board. Moreover they felt that the combination of speech and traces of hand movements gives a touch of sensibility and makes it easier to connect with the message. Participants also liked the naturalness and simplicity of the interaction. Future work includes implementing a fully automated system for segmentation.

9. REFERENCES

[1] Baudel, T. and Beaudouin-Lafon, M. 1993. Charade: remote control of objects using free-hand gestures. *Communications 36*, 7 (1993), ACM Press, 28-35.

[2] Clark, H. H., & Brennan, S.E. Grounding in Communication. In L.B. Resnick, R.M. Levine, & S.D. Teasley (Eds.). Perspectives on socially shared cognition, (1991), 127-149.

[3] Elsayed, A. Machine-Mediated Communication: The Techology. In *Proc. ICALT'06*, IEEE.

[4] Garner, S., and McDonagh-Philp, D. Problem Interpretation and Resolution via Visual Stimuli: The Use of 'Mood Boards' in Design Education. *International Journal of Art & Design Education 20*, 1 (2001), 57-64.

[5] Geyer, W., Richter, H., Abowd, G. 2005. Toward a Smarter Meeting Record – Capture and Access of Meetings Revisited. *Multimedia Tools and Applications 27* (2005), Springer Science, 393-410.

[6] Ju, S. X., Black, M. J., Minneman, S., and Kimber, D. 1997. Analysis of Gesture and Action in Technical Talks for Video Indexing. In *Proc. CVPR '97*, IEEE (1997).

[7] Kirk, D. and Stanton Fraser, D. 2006. Comparing remote gesture technologies for supporting collaborative physical tasks. In *Proc. CHI '06*, ACM Press (2006), 1191-1200.

[8] Lamming, M. G. *Towards a Human Memory Prosthesis.* Technical Report #EPC-91-116 EPC-91-116. Rank Xerox EuroPARC, 1991.

[9] Lucero, A., Aliakseyeu, D., and Martens, J.B. Augmenting Mood Boards: Flexible and Intuitive Interaction in the Context of the Design Studio. In *Proc. TableTop 2007*, IEEE (2007), 147-154.

[10] Lucero, A. and Mattelmäki, T. 2007. Professional Probes: A Pleasurable Little Extra for the Participant's Work. In *Proc. IASTED HCI 2007*, Acta Press, 170-176.

[11] Prante, T., Röcker, C., Streitz, N.A., Stenzel, R., Magerkurth, C., van Alphen, D., and Plewe, D.A. Hello.Wall - Beyond Ambient Displays. In *Adj. Proc. UBICOMP 2003*, 277-278.

[12] Stifelman, L. *The Audio Notebook: Paper and Pen Interaction with Structured Speech.* Ph.D. dissertation, MIT Media Laboratory, 1997.

[13] Vogel, D. and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proc. UIST 2004*, ACM Press (2004), 137-146.

[14] Wang, M. Q. and Hirschberg, J. Automatic Classification of Intonational Phrase Boundaries. Computer, Speech, and Language, 6:175–196, 1992.